AMERICAN JOURNAL OF MATHEMATICS

FOUNDED BY THE JOHNS HOPKINS UNIVERSITY

EDITED BY

G. D. BIRKHOFF
HARVARD UNIVERSITY

H. WEYL
THE INSTITUTE FOR ADVANCED STUDY

F. D. MURNAGHAN
THE JOHNS HOPKINS UNIVERSITY

R. L. WILDER
UNIVERSITY OF MICHIGAN

A. WINTNER
THE JOHNS HOPKINS UNIVERSITY

WITH THE COÖPERATION OF

G. BIRKHOFF	R. P. AGNEW	V. G. GROVE
N. DUNFORD	C. CHEVALLEY	M. H. HEINS
E. P. LANE	G. A. HEDLUND	D. C. LEWIS
D. MONTGOMERY	S. B. MYERS	T. RADÓ
J. L. SYNGE	N. E. STEENROD	H. S. WALL

PUBLISHED UNDER THE JOINT AUSPICES OF
THE JOHNS HOPKINS UNIVERSITY
AND
THE AMERICAN MATHEMATICAL SOCIETY

VOLUME LXV

1943

THE JOHNS HOPKINS PRESS
BALTIMORE, MARYLAND
U. S. A.

Mathematics Library

JAN 26 1943

AMERICAN JOURNAL OF MATHEMATICS

FOUNDED BY THE JOHNS HOPKINS UNIVERSITY

EDITED BY

G. D. BIRKHOFF
HARVARD UNIVERSITY

H. WEYL
THE INSTITUTE FOR ADVANCED STUDY

F. D. MURNAGHAN
THE JOHNS HOPKINS UNIVERSITY

R. L. WILDER
UNIVERSITY OF MICHIGAN

A. WINTNER
THE JOHNS HOPKINS UNIVERSITY

WITH THE COÖPERATION OF

G. BIRKHOFF
N. DUNFORD
E. P. LANE
D. MONTGOMERY
J. L. SYNGE
R. P. AGNEW
C. CHEVALLEY
G. A. HEDLUND
S. B. MYERS
N. E. STEENROD
V. G. GROVE
M. H. HEINS
D. C. LEWIS
T. RADÓ
H. S. WALL

PUBLISHED UNDER THE JOINT AUSPICES OF
THE JOHNS HOPKINS UNIVERSITY
AND
THE AMERICAN MATHEMATICAL SOCIETY

Volume LXV, Number 1
JANUARY, 1943

THE JOHNS HOPKINS PRESS
BALTIMORE, MARYLAND
U. S. A.

CONTENTS

	PAGE
Symplectic geometry. By CARL LUDWIG SIEGEL, ...	1
The theory of simple rings. By EMIL ARTIN and GEORGE WHAPLES, .	87
On weakly compact subsets of a Banach space. By R. S. PHILLIPS, .	108
Continua of finite linear measure I. By SAMUEL EILENBERG and O. G. HARROLD, JR., ...	137
A function-theoretic identity. By D. C. SPENCER, ...	147
On functions of class 1 having the property of Darboux. By ISAIAH MAXIMOFF, ...	161
The calculus of non-contradiction. By GLENN EDWARD DEXTER, .	171
Splittings of a lattice. By PHILIP M. WHITMAN, ...	179

The AMERICAN JOURNAL OF MATHEMATICS will appear four times yearly.

The subscription price of the JOURNAL for the current volume is $7.50 (foreign postage 50 cents); single numbers $2.00.

A few complete sets of the JOURNAL remain on sale.

Papers intended for publication in the JOURNAL may be sent to any of the Editors.

Editorial communications may be sent to Professor F. D. MURNAGHAN at The Johns Hopkins University.

Subscriptions to the JOURNAL and all business communications should be sent to THE JOHNS HOPKINS PRESS, BALTIMORE, MARYLAND, U. S. A.

Entered as second-class matter at the Baltimore, Maryland, Postoffice, acceptance for mailing at special rate of postage provided for in Section 1103, Act of October 3, 1917, Authorized on July 3, 1918.

PRINTED IN THE UNITED STATES OF AMERICA
BY J. H. FURST COMPANY, BALTIMORE, MARYLAND

SYMPLECTIC GEOMETRY.*

By CARL LUDWIG SIEGEL.

I. INTRODUCTION.

1. Our present knowledge concerning functions of several complex variables $z_1, \cdots, z_m$ is much less far-reaching and complete than the classical theory in the case $m = 1$. If we want to proceed further, it seems reasonable to investigate, in the first place, a *special* class of analytic functions of m variables found by the following considerations:

Let R be the Riemann surface of an analytic function of a single variable. On account of the main theorem of uniformization, the universal covering surface U of R can be conformally mapped onto a simple domain E, which is either the unit-circle $|z| < 1$ or the finite z-plane or the complete z-plane. The conformal mappings of E onto itself form a group Ω of linear transformations, and the fundamental group of R is faithfully represented by a subgroup Δ of Ω, discontinuous in E. By the introduction of the uniformizing parameter z, the general theory of the analytic single-valued functions on R is reduced to the theory of the automorphic functions with the group Δ.

The group Ω is transitive, i. e., there exists for any two points z_1 and z_2 of E an element of Ω transforming z_1 into z_2. Moreover, there exists for every point z_1 of E an involution in Ω with the fixed point z_1, i. e., an element of Ω identical with its inverse and transforming z_1 into itself. Consequently E is a *sym.. etric* space, in the notation of Elie Cartan. The domain E is bounded, if we consider only the first case, the case of the unit-circle $|z| < 1$; it is well known, that this occurs if and only if U has at least two frontier points.

A generalization of the theory of automorphic functions to the case of an arbitrary number of variables requires the following three steps: 1) To determine all bounded simple domains E in the space of m complex variables, which are symmetric spaces with respect to a group Ω of analytic mappings. 2) To investigate the invariant geometric properties of E, to find the discontinuous subgroups Δ of Ω and to construct their fundamental domains. 3) To study the field of automorphic functions in E with the group Δ.

The first step has been made by Cartan; he obtained explicitly 6 different types of irreducible domains E, such that all other bounded simple symmetric

* Received February 27, 1942.

analytic spaces can be derived from them by analytic mappings and topological products.

We shall consider more closely the second step. We restrict our researches to one of the six possible types, which is the most important for applications to other branches of mathematics. In this case, the number m of complex dimensions is $\frac{1}{2}n(n+1)$, with integral $n \geq 1$, the m variables form the elements $z_{kl} = z_{lk}$ $(1 \leq k \leq l \leq n)$ of a symmetric complex matrix $\mathfrak{Z} = (z_{kl})$ with n rows and E consists of all points $\mathfrak{Z}$ for which the hermitian form $\sum_{k=1}^{n} (|u_k|^2 - \sum_{l=1}^{n} |z_{kl}u_l|^2)$ in the auxiliary variables $u_1, \cdots, u_n$ is positive definite.

The third step has already been carried out, in a former publication, for the special case of the modular group of degree n. It is possible to generalize most of those results, but we shall not do so in the present paper.

2. Notations, definitions, results. All German letters denote matrices with complex elements; small German letters denote columns. The upper indices p and q in $\mathfrak{A}^{(pq)}$ designate the number p of rows and the number q of columns of the matrix $\mathfrak{A}$; instead of $\mathfrak{A}^{(pp)}$ and $\mathfrak{a}^{(p1)}$ we write more simply $\mathfrak{A}^{(p)}$ and $\mathfrak{a}^{(p)}$. If $a_1, \cdots, a_p$ are the diagonal elements of $\mathfrak{A}^{(p)}$ and if all other elements are 0, we write $\mathfrak{A} = [a_1, \cdots, a_p]$ and call $\mathfrak{A}$ a *diagonal* matrix. The letter $\mathfrak{E}$ denotes a *unit* matrix, and 0 denotes also a *zero* matrix. If $\mathfrak{B}$ is any matrix, $\mathfrak{B}'$ is the *transposed* matrix and $\bar{\mathfrak{B}}$ the *conjugate complex* matrix. We use the abbreviations $\mathfrak{B}'\mathfrak{A}\mathfrak{B} = \mathfrak{A}[\mathfrak{B}]$, $\mathfrak{B}'\mathfrak{A}\bar{\mathfrak{B}} = \mathfrak{A}\{\mathfrak{B}\}$. The inequality $\mathfrak{A} > 0$ means that $\mathfrak{A} = \bar{\mathfrak{A}}'$ is the matrix of a positive definite hermitian form, i. e., $\mathfrak{A}\{\mathfrak{x}\} > 0$ for all $\mathfrak{x} \neq 0$; obviously $\mathfrak{A} > 0$ means in the case of a real $\mathfrak{A}$, that $\mathfrak{A} = \mathfrak{A}'$ is the matrix of a positive definite quadratic form, i. e., $\mathfrak{A}[\mathfrak{x}] > 0$ for all real $\mathfrak{x} \neq 0$. The *trace* $\sigma(\mathfrak{A})$ of a matrix $\mathfrak{A}^{(p)} = (a_{kl})$ is defined by $\sigma(\mathfrak{A}) = \sum_{k=1}^{p} a_{kk}$.

We denote by $\mathfrak{Z} = (z_{kl})$ a symmetric matrix with n rows and variable complex elements $z_{kl} = z_{lk}$ $(1 \leq k \leq l \leq n)$; $\mathfrak{Z} = \mathfrak{X} + i\mathfrak{Y}$, where $\mathfrak{X} = \frac{1}{2}(\mathfrak{Z} + \bar{\mathfrak{Z}})$ and $\mathfrak{Y} = (1/2i)(\mathfrak{Z} - \bar{\mathfrak{Z}})$ are the real and imaginary part of $\mathfrak{Z}$. The condition $\mathfrak{E} - \mathfrak{Z}\bar{\mathfrak{Z}} > 0$ defines a bounded domain E which is obviously a generalization of the unit-circle. On the other hand, the domain H defined by the inequality $\mathfrak{Y} > 0$ is a generalization of the upper half-plane. It is well-known that the transformation $w = (az + b)/(cz + d)$ with real a, b, c, d and $ad - bc = 1$ is the most general analytic mapping of the upper half-plane onto itself. In order to generalize this theorem, we have to introduce the symplectic group. The *homogeneous symplectic* group Ω_0 consists of all real matrices

$$\mathfrak{M} = \begin{pmatrix} \mathfrak{A}^{(n)} & \mathfrak{B}^{(n)} \\ \mathfrak{C}^{(n)} & \mathfrak{D}^{(n)} \end{pmatrix}$$

satisfying the condition $\mathfrak{J}[\mathfrak{M}] = \mathfrak{J}$ with

$$\mathfrak{J} = \begin{pmatrix} 0 & \mathfrak{E}^{(n)} \\ -\mathfrak{E}^{(n)} & 0 \end{pmatrix}.$$

It is easily proved that the transformation

$$\mathfrak{W} = (\mathfrak{A}\mathfrak{Z} + \mathfrak{B})(\mathfrak{C}\mathfrak{Z} + \mathfrak{D})^{-1} \tag{1}$$

maps the domain H onto itself. These transformations form the (inhomogeneous) *symplectic* group Ω obtained by identifying $\mathfrak{M}$ and $-\mathfrak{M}$.

THEOREM 1. *Every analytic mapping of H onto itself is symplectic.*

The next four theorems generalize known properties of the Poincaré model of non-euclidean geometry. For any two points $\mathfrak{Z}, \mathfrak{Z}_1$ of H we define

$$\mathfrak{R}(\mathfrak{Z}, \mathfrak{Z}_1) = (\mathfrak{Z} - \mathfrak{Z}_1)(\mathfrak{Z} - \bar{\mathfrak{Z}}_1)^{-1}(\bar{\mathfrak{Z}} - \bar{\mathfrak{Z}}_1)(\bar{\mathfrak{Z}} - \mathfrak{Z}_1)^{-1}.$$

THEOREM 2. *There exists a symplectic transformation mapping a given pair $\mathfrak{Z}, \mathfrak{Z}_1$ of H into another given pair $\mathfrak{W}, \mathfrak{W}_1$ of H, if and only if the two matrices $\mathfrak{R}(\mathfrak{Z}, \mathfrak{Z}_1)$ and $\mathfrak{R}(\mathfrak{W}, \mathfrak{W}_1)$ have the same characteristic roots.*

Let $d\mathfrak{Z} = (dz_{kl})$ denote the matrix of the differentials dz_{kl}. The quadratic differential form

$$ds^2 = \sigma(\mathfrak{Y}^{-1} d\mathfrak{Z} \mathfrak{Y}^{-1} d\bar{\mathfrak{Z}}) \tag{2}$$

is invariant under Ω and defines a Riemann metric in H.

THEOREM 3. *There exists exactly one geodesic arc connecting two arbitrary points $\mathfrak{Z}, \mathfrak{Z}_1$ of H; its length ρ is given by*

$$\rho^2 = \sigma\left(\log^2 \frac{1 + \mathfrak{R}^{\frac{1}{2}}}{1 - \mathfrak{R}^{\frac{1}{2}}}\right)$$

with $\mathfrak{R} = \mathfrak{R}(\mathfrak{Z}, \mathfrak{Z}_1)$ and

$$\log^2 \frac{1 + \mathfrak{R}^{\frac{1}{2}}}{1 - \mathfrak{R}^{\frac{1}{2}}} = 4\mathfrak{R}\left(\sum_{k=0}^{\infty} \frac{\mathfrak{R}^k}{2k+1}\right)^2.$$

THEOREM 4. *All geodesics are symplectic images of the curves $\mathfrak{Z} = i[p_1^s, \cdots, p_n^s]$, where $p_1, \cdots, p_n$ are arbitrary positive constants satisfying $\sum_{k=1}^{n} \log^2 p_k = 1$.*

Let $\mathfrak{X} = (x_{kl})$, $\mathfrak{Y}^{-1} = (Y_{kl})$ and dv be the euclidean volume element in the space with the $n(n+1)$ rectangular cartesian coördinates x_{kl}, Y_{kl}

$(1 \leq k \leq l \leq n)$. It is easily shown that $2^{n(n-1)/2}dv$ is the volume element for the symplectic metric (2).

THEOREM 5. *The Euler characteristic of a closed manifold F with the metric (2) is*

$$\chi = c_n(-\pi)^{-n(n+1)/2}\int_F dv, \tag{3}$$

where c_n denotes a positive rational number depending only upon n; in particular $c_1 = \frac{1}{2}$, $c_2 = \frac{3}{8}$, $c_3 = \frac{45}{256}$.

The theorems 6 and 7 are concerned with the generalization of the Fuchsian groups and their fundamental domains. Let Δ be a subgroup of the symplectic group Ω. Two points $\mathfrak{Z}, \mathfrak{W}$ of H are called *equivalent* under Δ, if (1) holds for a matrix $\mathfrak{M}$ of Δ. The group Δ is *discontinuous*, if no set of equivalent points has a limit point in H. A domain F in H is a *fundamental domain* for Δ, if the images of F under Δ cover H without gaps and overlappings. A domain F is called a *star*, if there exists an inner point $\mathfrak{Z}_0$ of F such that for every point $\mathfrak{Z}$ of F the whole geodesic arc between $\mathfrak{Z}$ and $\mathfrak{Z}_0$ belongs to F.

THEOREM 6. *A fundamental domain F of a discontinuous group Δ may be chosen as a star bounded by analytic surfaces and such that every compact domain in H is covered by a finite number of images of F under Δ.*

A discontinuous group Δ is called of the *first kind*, if there exists a *normal* fundamental domain F having the following three properties: 1) Every compact domain in H is covered by a finite number of images of F; 2) only a finite number of images of F are neighbors of F; 3) the integral

$$V(\Delta) = \int_F dv$$

converges. The space H is called *compact relative to* Δ, if there exists for every infinite sequence of points $\mathfrak{Z}_k$ $(k = 1, 2, 3, \cdots)$ in H a compact sequence $\mathfrak{W}_k$ such that $\mathfrak{W}_k$ is equivalent to $\mathfrak{Z}_k$ under Δ.

THEOREM 7. *If H is compact relative to a discontinuous group Δ, then Δ is of the first kind and has a compact normal fundamental domain.*

Let us assume that a discontinuous group Δ has no fixed point in H, i. e., that no transformation of Δ except the identical one has a fixed point in H. Identifying equivalent frontier points of a fundamental domain F, we obtain a closed manifold, if H is compact relative to Δ. The Euler number of

this manifold is then given by Theorem 5. If H is not compact relative to Δ, we obtain an open manifold. It is probable that Theorem 5 still holds good for this open manifold, provided Δ is of the first kind; it may easily be proved that this is true in the case $n = 1$.

The rest of the paper deals with two special classes of groups Δ defined by arithmetical properties. The simplest and most important example of a discontinuous subgroup of Ω is given by the *modular* group Γ of degree n consisting of all symplectic matrices $\mathfrak{M}^{(2n)}$ with rational integral elements.

THEOREM 8. *The modular group of degree n is a discontinuous group of the first kind.*

Let K be a totally real algebraic number-field of degree $h \geq 1$, $K(\sqrt{-r})$ a totally imaginary quadratic field over K and s a positive number in K such that all other conjugates of s are negative. Let $\mathfrak{G}^{(2n)}$ be a skew-symmetric matrix and $\mathfrak{H}^{(2n)}$ a hermitian matrix, both with elements from $K(\sqrt{-r})$ and non-singular. We assume that all conjugates of $\mathfrak{H}$ except $\mathfrak{H}$ and $\bar{\mathfrak{H}}$ are positive and that $\mathfrak{G}$ and $\mathfrak{H}$ are connected by the relation $\mathfrak{H}\bar{\mathfrak{G}}^{-1}\bar{\mathfrak{H}} = s\mathfrak{G}$. Let $\Lambda(\mathfrak{G}, \mathfrak{H})$ denote the group of matrices $\mathfrak{U}$ with integral elements of $K(\sqrt{-r})$ satisfying the two conditions $\mathfrak{G}[\mathfrak{U}] = \mathfrak{G}$ and $\mathfrak{H}\{\mathfrak{U}\} = \mathfrak{H}$. Then there exists a constant matrix $\mathfrak{C}$ such that $\mathfrak{C}^{-1}\mathfrak{U}\mathfrak{C} = \mathfrak{M}$ is symplectic, and $\mathfrak{C}^{-1}\Lambda(\mathfrak{G}, \mathfrak{H})\mathfrak{C} = \Lambda(\mathfrak{G}, \mathfrak{H})$ is a subgroup of Ω. The modular group Γ is a particular case of these groups $\Delta(\mathfrak{G}, \mathfrak{H})$, namely the case $h = 1$, $\mathfrak{G} = \mathfrak{J}$, $\mathfrak{H} = i\mathfrak{J}$, $r = 1$.

THEOREM 9. *The group $\Delta(\mathfrak{G}, \mathfrak{H})$ is discontinuous and of the first kind. In the case $h > 1$, the space H is compact relative to $\Delta(\mathfrak{G}, \mathfrak{H})$.*

For every ideal κ of $K(\sqrt{-r})$, we denote by $\Lambda_\kappa(\mathfrak{G}, \mathfrak{H})$ the congruence subgroup of $\Lambda(\mathfrak{G}, \mathfrak{H})$ defined by the condition $\mathfrak{U} \equiv \mathfrak{E} \pmod{\kappa}$, and by $\Delta_\kappa(\mathfrak{G}, \mathfrak{H}) = \mathfrak{C}^{-1}\Lambda_\kappa(\mathfrak{G}, \mathfrak{H})\mathfrak{C}$ the corresponding subgroup of $\Delta(\mathfrak{G}, \mathfrak{H})$.

THEOREM 10. *Let ρ be a prime ideal of $K(\sqrt{-r})$ and κ the least power of ρ such that p is not divisible by κ^{p-1}, where p denotes the rational prime number divisible by ρ. Then $\Delta_\kappa(\mathfrak{G}, \mathfrak{H})$ has no fixed point in H.*

On account of the theorems 5 and 10, the calculation of the integral $V(\Delta)$ for $\Delta = \Delta(\mathfrak{G}, \mathfrak{H})$ is important. Applying the Gauss-Dirichlet method from analytic number theory, we obtain in the case of the modular group of degree n a curious connection with Riemann's ζ-function. Using the abbreviation $\xi(t) = \pi^{-(t/2)}\Gamma(t/2)\zeta(t)$, so that $\xi(t) = \xi(1-t)$ is the functional equation of $\zeta(t)$, we have

THEOREM 11. *The symplectic volume of the fundamental domain of the modular group is*

$$V(\Gamma) = 2\prod_{k=1}^{n}\xi(2k).$$

This formula may be written in a different way, suggested by the results of the analytic theory of quadratic forms. Consider a domain Q in the space of the real skew-symmetric matrices $\mathfrak{Q}^{(2n)} = (q_{kl})$, with the rectangular cartesian coördinates q_{kl} $(1 \leq k < l \leq 2n)$, and denote by L the corresponding part of the space of the real matrices $\mathfrak{L}^{(2n)}$ defined by the condition $\mathfrak{J}[\mathfrak{L}] = \mathfrak{Q}$, the coördinates in L being the $4n^2$ elements of $\mathfrak{L}$. Obviously L is invariant under any mapping $\mathfrak{L} \to \mathfrak{M}\mathfrak{L}$ with symplectic $\mathfrak{M}$. Let L_0 be a fundamental domain in L with respect to the homogeneous modular group, and let $v(L_0)$, $v(Q)$ be the euclidean volumes of L_0 and Q. We define

$$d_0(\Gamma) = \lim_{Q \to \mathfrak{J}} \frac{v(L_0)}{v(Q)}, \tag{4}$$

where Q runs over a sequence of domains tending to the single point $\mathfrak{J}$. On the other hand, let p be a rational prime number and E_p the number of modulo p incongruent integral solutions $\mathfrak{M}$ of the congruence $\mathfrak{J}[\mathfrak{M}] \equiv \mathfrak{J} \pmod{p}$. Since there are, modulo p, exactly $p^{n(2n-1)}$ integral skew-symmetric matrices $\mathfrak{Q}$ and p^{4n^2} integral matrices $\mathfrak{L}$, the expression

$$d_p(\Gamma) = p^{n(2n+1)}E_p^{-1}$$

may be considered as the p-adic analogue of $d_0(\Gamma)$. As a consequence of Theorem 11, we obtain

THEOREM 12. *Let p run over all prime numbers, then*

$$d_0(\Gamma) = \prod_p d_p(\Gamma).$$

It is possible to generalize this theorem for the case of an arbitrary group $\Delta(\mathfrak{G}, \mathfrak{H})$ instead of Γ.

Two subgroups Δ and Δ_1 of Ω are *conjugate*, if the relation $\Delta_1 = \mathfrak{F}^{-1}\Delta\mathfrak{F}$ holds for a symplectic matrix $\mathfrak{F}$. More generally, Δ and Δ_1 are called *commensurable*, if they contain conjugate subgroups of finite index. It is important, for the theory of automorphic functions, to know whether two given groups Δ and Δ_1 are commensurable or not. Let $\Delta = \Delta(\mathfrak{G}, \mathfrak{H})$, $\Delta_1 = \Delta(\mathfrak{G}_1, \mathfrak{H}_1)$ and let K_1, r_1, s_1 have the same meaning for $\mathfrak{G}_1, \mathfrak{H}_1$ that K, r, s have for $\mathfrak{G}, \mathfrak{H}$.

THEOREM 13. *The two groups $\Delta(\mathfrak{G}, \mathfrak{H})$ and $\Delta(\mathfrak{G}_1, \mathfrak{H}_1)$ are commensurable if, and only if, $K = K_1$ and the ternary quadratic forms $rsx^2 - ry^2 + sz^2$ and $r_1s_1x^2 - r_1y^2 + s_1z^2$ are equivalent in K.*

In the particular case $n = 2$, another class of discontinuous subgroups of Ω is given by the theory of *units* of quinary quadratic forms. Let K be again a totally real field of degree h. We consider a quadratic form $\mathfrak{T}[\mathfrak{x}]$ of 5 variables with coefficients from K and assume that $\mathfrak{T}[\mathfrak{x}]$ has the signature 2, 3, whereas all other conjugates of $\mathfrak{T}[\mathfrak{x}]$ are definite. Let $\Lambda(\mathfrak{T})$ be the group of all integral matrices $\mathfrak{U}$ in K satisfying $\mathfrak{T}[\mathfrak{U}] = \mathfrak{T}$, $|\mathfrak{U}| = 1$. On account of the spin representation of the orthogonal group, either $\Lambda(\mathfrak{T})$ itself or a subgroup of index 2 is then isomorphic to a certain subgroup $\Delta(\mathfrak{T})$ of Ω. Concerning these groups $\Delta(\mathfrak{T})$, there are analogues of the theorems 9, 10, 12, 13; in particular, analogous to Theorem 9, we have

THEOREM 14. *The group $\Delta(\mathfrak{T})$ is discontinuous and of the first kind. In the case $h > 1$, the space H is compact relative to $\Delta(\mathfrak{T})$.*

It would be interesting to seek discontinuous subgroups of Ω which are not commensurable with any of the groups $\Delta(\mathfrak{G}, \mathfrak{H})$ and $\Delta(\mathfrak{T})$. In the case $n = 1$, we may start with an arbitrary polygon satisfying certain conditions, and use the reflection method, but this simple geometric principle breaks down for $n > 1$.

---

3. Literature.

C. B. Allendoerfer, "The Euler number of a Riemann manifold," *American Journal of Mathematics* **62**, pp. 243-248 (1940).

E. Cartan, "Sur les domaines bornés homogènes de l'espace de n variables complexes," *Abhandlungen aus dem Mathematischen Seminar der Hansischen Universität* **11**, pp. 116-162 (1936).

W. Fenchel, "On total curvatures of Riemannian manifolds: I," *The Journal of the London Mathematical Society* **15**, pp. 15-22 (1940).

R. Fricke and F. Klein, *Vorlesungen über die Theorie der automorphen Funktionen*, Vol. 1, B. G. Teubner, Leipzig (1897).

G. Fubini, "A remark on general Fuchsian groups," *Proceedings of the National Academy of Sciences* **26**, pp. 695-700 (1940).

———, "The distance in general Fuchsian geometries," *Proceedings of the National Academy of Sciences* **26**, pp. 700-708 (1940).

G. Giraud, "Sur une classe de groupes discontinus de transformations birationnelles quadratiques et sur les fonctions de trois variables indépendantes restant invariables par ces transformations," *Annales Scientifiques de l'Ecole Normale Supérieure* (3) **32**, pp. 237-403 (1915).

———, "Sur les groupes de transformations semblables arithmétiques de certaines formes quadratiques quinaires indéfinies et sur les fonctions de trois variables indépendantes invariantes par des groupes isomorphes aux précédents," *Annales Scientifiques de l'Ecole Normale Supérieure* (3) **33**, pp. 330-362 (1916).

P. Humbert, "Théorie de la réduction des formes quadratiques définies positives dans un corps algébrique K fini," *Commentarii Mathematici Helvetici* **12**, pp. 263-306 (1940).

H. Minkowski, "Diskontinutätsbereich für arithmetische Äquivalenz," *Journal für die reine und angewandte Mathematik* **129**, pp. 220-274 (1905).

C. L. Siegel, "Einführung in die Theorie der Modulfunktionen n-ten Grades," *Mathematische Annalen* **116**, pp. 617-657 (1939).

———, "Einheiten quadratischer Formen," *Abhandlungen aus dem Mathematischen Seminar der Hansischen Universität* **13**, pp. 209-239 (1940).

M. Sugawara, "Über eine allgemeine Theorie der Fuchsschen Gruppen und Theta-Reihen," *Annals of Mathematics* (2) **41**, pp. 488-494 (1940).

———, "A generalization of Poincaré-space," *Proceedings of the Imperial Academy of Japan* **16**, pp. 373-377 (1940).

E. Witt, "Eine Identität zwischen Modulformen zweiten Grades," *Abhandlungen aus dem Mathematischen Seminar der Hansischen Universität* **14**, pp. 323-337 (1941).

II. THE SYMPLECTIC GROUP.

4. The linear substitution $z = i\dfrac{1+z_0}{1-z_0}$ maps the unit-circle $z_0\bar{z}_0 < 1$ onto the upper half-plane $\dfrac{1}{2i}(z-\bar{z}) > 0$. We shall prove that there is an immediate generalization to the case $n > 1$.

Let $\mathfrak{Z}_0$ be a point of E, i. e., $\mathfrak{Z}_0 = \mathfrak{Z}_0'$, $\mathfrak{E} - \mathfrak{Z}_0\bar{\mathfrak{Z}}_0 > 0$. If $\mathfrak{x}$ is a solution of $(\mathfrak{E} - \mathfrak{Z}_0)\mathfrak{x} = 0$, then $\bar{\mathfrak{x}} = \bar{\mathfrak{Z}}_0\bar{\mathfrak{x}}$, $\mathfrak{x}' = \mathfrak{x}'\mathfrak{Z}_0$ and consequently $(\mathfrak{E} - \mathfrak{Z}_0\bar{\mathfrak{Z}}_0)\{\mathfrak{x}\} = \mathfrak{x}'\bar{\mathfrak{x}} - \mathfrak{x}'\mathfrak{Z}_0\bar{\mathfrak{Z}}_0\bar{\mathfrak{x}} = 0$, $\mathfrak{x} = 0$. This proves $|\mathfrak{E} - \mathfrak{Z}_0| \neq 0$ and the existence of the matrix

$$i(\mathfrak{E} + \mathfrak{Z}_0)(\mathfrak{E} - \mathfrak{Z}_0)^{-1} = \mathfrak{Z}. \tag{5}$$

Obviously $\mathfrak{Z} = \mathfrak{Z}'$ and

$$\frac{1}{2i}(\mathfrak{Z} - \bar{\mathfrak{Z}}) = \tfrac{1}{2}(\mathfrak{E} - \mathfrak{Z}_0)^{-1}\left((\mathfrak{E} + \mathfrak{Z}_0)(\mathfrak{E} - \bar{\mathfrak{Z}}_0) + (\mathfrak{E} - \mathfrak{Z}_0)(\mathfrak{E} + \bar{\mathfrak{Z}}_0)\right)(\mathfrak{E} - \bar{\mathfrak{Z}}_0)^{-1}$$
$$= (\mathfrak{E} - \mathfrak{Z}_0\bar{\mathfrak{Z}}_0)\{(\mathfrak{E} - \mathfrak{Z}_0)^{-1}\} > 0;$$

hence $\mathfrak{Z}$ is a point of H. On the other hand, let $\mathfrak{Z}$ be an arbitrary point of H, i. e., $\mathfrak{Z} = \mathfrak{Z}'$, $\dfrac{1}{2i}(\mathfrak{Z} - \bar{\mathfrak{Z}}) > 0$. If $\mathfrak{x}$ is a solution of $(\mathfrak{Z} + i\mathfrak{E})\mathfrak{x} = 0$, then $\bar{\mathfrak{Z}}\bar{\mathfrak{x}} = i\bar{\mathfrak{x}}$, $\mathfrak{x}'\mathfrak{Z} = -i\mathfrak{x}'$ and consequently $\dfrac{1}{2i}(\mathfrak{Z} - \bar{\mathfrak{Z}})\{\mathfrak{x}\} = \dfrac{1}{2i}(\mathfrak{x}'\mathfrak{Z}\bar{\mathfrak{x}} - \mathfrak{x}'\bar{\mathfrak{Z}}\bar{\mathfrak{x}})$ $= -\mathfrak{x}'\bar{\mathfrak{x}} \leq 0$, $\mathfrak{x} = 0$. This proves $|\mathfrak{Z} + i\mathfrak{E}| \neq 0$ and the existence of the matrix

$$(\mathfrak{Z} - i\mathfrak{E})(\mathfrak{Z} + i\mathfrak{E})^{-1} = \mathfrak{Z}_0. \tag{6}$$

Obviously $\mathfrak{Z}_0 = \mathfrak{Z}'_0$ and

$$\mathfrak{E} - \mathfrak{Z}_0\bar{\mathfrak{Z}}_0 = (\mathfrak{Z} + i\mathfrak{E})^{-1}((\mathfrak{Z} + i\mathfrak{E})(\bar{\mathfrak{Z}} - i\mathfrak{E}) - (\mathfrak{Z} - i\mathfrak{E})(\bar{\mathfrak{Z}} + i\mathfrak{E}))(\bar{\mathfrak{Z}} - i\mathfrak{E})^{-1}$$
$$= \frac{2}{i}(\mathfrak{Z} - \bar{\mathfrak{Z}})\{(\mathfrak{Z} + i\mathfrak{E})^{-1}\} > 0;$$

hence $\mathfrak{Z}_0$ is a point of E. Moreover (6) follows from (5), and vice versa.

5. The homogeneous symplectic group Ω_0 consists of all matrices

$$\mathfrak{M} = \begin{pmatrix} \mathfrak{A} & \mathfrak{B} \\ \mathfrak{C} & \mathfrak{D} \end{pmatrix}$$

with real elements satisfying $\mathfrak{M}'\mathfrak{J}\mathfrak{M} = \mathfrak{J}$. Since $\mathfrak{J}^{-1} = -\mathfrak{J}$, we have then also $\mathfrak{M}\mathfrak{J}\mathfrak{M}' = \mathfrak{J}$; hence $\mathfrak{M}'$ is symplectic and

$$\mathfrak{A}\mathfrak{B}' = \mathfrak{B}\mathfrak{A}', \qquad \mathfrak{C}\mathfrak{D}' = \mathfrak{D}\mathfrak{C}', \qquad \mathfrak{A}\mathfrak{D}' - \mathfrak{B}\mathfrak{C}' = \mathfrak{E}. \tag{7}$$

Let $\mathfrak{Z}$ be a point of H, i. e.,

$$\mathfrak{J}\begin{bmatrix} \mathfrak{Z} \\ \mathfrak{E} \end{bmatrix} = 0, \qquad \frac{1}{2i}\mathfrak{J}\begin{Bmatrix} \mathfrak{Z} \\ \mathfrak{E} \end{Bmatrix} > 0.$$

The matrices $\mathfrak{A}\mathfrak{Z} + \mathfrak{B} = \mathfrak{P}$, $\mathfrak{C}\mathfrak{Z} + \mathfrak{D} = \mathfrak{Q}$ satisfy

$$\mathfrak{M}\begin{pmatrix} \mathfrak{Z} \\ \mathfrak{E} \end{pmatrix} = \begin{pmatrix} \mathfrak{P} \\ \mathfrak{Q} \end{pmatrix}, \qquad \mathfrak{J}\begin{bmatrix} \mathfrak{P} \\ \mathfrak{Q} \end{bmatrix} = 0, \qquad \frac{1}{2i}\mathfrak{J}\begin{Bmatrix} \mathfrak{P} \\ \mathfrak{Q} \end{Bmatrix} > 0,$$

$$\mathfrak{P}'\mathfrak{Q} = \mathfrak{Q}'\mathfrak{P}, \qquad \frac{1}{2i}(\mathfrak{P}'\bar{\mathfrak{Q}} - \mathfrak{Q}'\bar{\mathfrak{P}}) > 0.$$

If $\mathfrak{x}$ is a solution of $\mathfrak{Q}\mathfrak{x} = 0$, then $\bar{\mathfrak{Q}}\bar{\mathfrak{x}} = 0$, $\mathfrak{x}'\mathfrak{Q}' = 0$, $\frac{1}{2i}(\mathfrak{P}'\bar{\mathfrak{Q}} - \mathfrak{Q}'\bar{\mathfrak{P}})\{\mathfrak{x}\} = 0$, whence $\mathfrak{x} = 0$, $|\mathfrak{Q}| \neq 0$. This proves the existence of

$$(\mathfrak{A}\mathfrak{Z} + \mathfrak{B})(\mathfrak{C}\mathfrak{Z} + \mathfrak{D})^{-1} = \mathfrak{P}\mathfrak{Q}^{-1} = \mathfrak{W},$$

with $\mathfrak{W} = \mathfrak{W}'$, $\frac{1}{2i}(\mathfrak{W} - \bar{\mathfrak{W}})\{\mathfrak{Q}\} > 0$, $\frac{1}{2i}(\mathfrak{W} - \bar{\mathfrak{W}}) > 0$. Consequently the fractional linear transformation

$$\mathfrak{W} = (\mathfrak{A}\mathfrak{Z} + \mathfrak{B})(\mathfrak{C}\mathfrak{Z} + \mathfrak{D})^{-1} = (\mathfrak{Z}\mathfrak{C}' + \mathfrak{D}')^{-1}(\mathfrak{Z}\mathfrak{A}' + \mathfrak{B}')$$

maps H into itself. Since $\mathfrak{Z}(-\mathfrak{C}'\mathfrak{W} + \mathfrak{A}') = \mathfrak{D}'\mathfrak{W} - \mathfrak{B}'$ and

$$\mathfrak{M}^{-1} = \mathfrak{J}^{-1}\mathfrak{M}'\mathfrak{J} = \begin{pmatrix} \mathfrak{D}' & -\mathfrak{B}' \\ -\mathfrak{C}' & \mathfrak{A}' \end{pmatrix}, \tag{8}$$

we obtain $|-\mathfrak{C}'\mathfrak{W} + \mathfrak{A}'| \neq 0$ and $\mathfrak{Z} = (\mathfrak{D}'\mathfrak{W} - \mathfrak{B}')(-\mathfrak{C}'\mathfrak{W} + \mathfrak{A}')^{-1}$; hence H is mapped onto itself.

It is easily seen that two symplectic matrices $\mathfrak{M}$ and $\mathfrak{M}_1$ define the same symplectic mapping $\mathfrak{W} = (\mathfrak{A}\mathfrak{Z} + \mathfrak{B})(\mathfrak{C}\mathfrak{Z} + \mathfrak{D})^{-1}$, if and only if $\mathfrak{M}_1 = \pm \mathfrak{M}$. Hence the inhomogeneous symplectic group Ω is the factor group of Ω_0 obtained by identifying $\mathfrak{M}$ and $-\mathfrak{M}$.

We have

$$\begin{pmatrix} i\mathfrak{C}' + \mathfrak{D}' & -i\mathfrak{A}' - \mathfrak{B}' \\ 0 & \mathfrak{E} \end{pmatrix} \mathfrak{M} \begin{pmatrix} \mathfrak{E} & i\mathfrak{E} \\ 0 & \mathfrak{E} \end{pmatrix} = \begin{pmatrix} \mathfrak{E} & 0 \\ \mathfrak{C} & i\mathfrak{C} + \mathfrak{D} \end{pmatrix}.$$

On the other hand $|\, i\mathfrak{C} + \mathfrak{D} \,| \neq 0$, since $i\mathfrak{E}$ is a point of H. This proves $|\, \mathfrak{M} \,| = 1$.

6. The fractional linear transformation (5) maps E onto H; its matrix is

$$\mathfrak{L} = \begin{pmatrix} i\mathfrak{E} & i\mathfrak{E} \\ -\mathfrak{E} & \mathfrak{E} \end{pmatrix}$$

and satisfies

$$\mathfrak{J}[\mathfrak{L}] = 2i\mathfrak{J}, \qquad \frac{1}{i}\,\mathfrak{J}\{\mathfrak{L}\} = 2\mathfrak{K}$$

with

$$\mathfrak{K} = \begin{pmatrix} -\mathfrak{E} & 0 \\ 0 & \mathfrak{E} \end{pmatrix}.$$

Let $\mathfrak{M}$ be an arbitrary symplectic matrix, i. e., $\mathfrak{J}[\mathfrak{M}] = \mathfrak{J}$, $\mathfrak{J}\{\mathfrak{M}\} = \mathfrak{J}$. Then

$$\mathfrak{L}^{-1}\mathfrak{M}\mathfrak{L} = \mathfrak{M}_0 = \begin{pmatrix} \mathfrak{A}_0 & \mathfrak{B}_0 \\ \mathfrak{C}_0 & \mathfrak{D}_0 \end{pmatrix}$$

fulfills the conditions $\mathfrak{J}[\mathfrak{M}_0] = \mathfrak{J}$, $\mathfrak{K}\{\mathfrak{M}_0\} = \mathfrak{K}$, whence $\mathfrak{J}[\mathfrak{M}_0] = \mathfrak{J}$, $\mathfrak{F}\bar{\mathfrak{M}}_0 = \mathfrak{M}_0\mathfrak{F}$ with

$$\mathfrak{F} = \mathfrak{J}\mathfrak{K} = \begin{pmatrix} 0 & \mathfrak{E} \\ \mathfrak{E} & 0 \end{pmatrix},$$

or more explicitly

$$(9) \qquad \mathfrak{A}_0\mathfrak{B}'_0 = \mathfrak{B}_0\mathfrak{A}'_0, \quad \mathfrak{A}_0\bar{\mathfrak{A}}'_0 - \mathfrak{B}_0\bar{\mathfrak{B}}'_0 = \mathfrak{E}, \quad \mathfrak{C}_0 = \bar{\mathfrak{B}}_0, \quad \mathfrak{D}_0 = \bar{\mathfrak{A}}_0.$$

The corresponding transformation

$$(10) \qquad \mathfrak{W}_0 = (\mathfrak{A}_0\mathfrak{Z}_0 + \mathfrak{B}_0)(\bar{\mathfrak{B}}_0\mathfrak{Z}_0 + \bar{\mathfrak{A}}_0)^{-1}$$

maps E onto itself, and all these transformations form the group $\mathfrak{L}^{-1}\Omega\mathfrak{L} = \Omega_E$.

We shall prove that Ω_E is transitive. Let $\mathfrak{Z}_0$ be any point of E, i. e., $\mathfrak{Z}_0 = \mathfrak{Z}'_0$, $\mathfrak{E} - \mathfrak{Z}_0\bar{\mathfrak{Z}}_0 > 0$. It is sufficient to prove the existence of a transformation (10) mapping $\mathfrak{Z}_0$ into 0. We choose $\mathfrak{A}_0$ such that $\mathfrak{A}_0(\mathfrak{E} - \mathfrak{Z}_0\bar{\mathfrak{Z}}_0)\bar{\mathfrak{A}}' = \mathfrak{E}$ and define $\mathfrak{B}_0 = -\mathfrak{A}_0\mathfrak{Z}_0$; then (9) is satisfied and (10) has obviously the required property. Consequently Ω is also transitive.

A mapping of Ω_E has the fixed point 0, if and only if $\mathfrak{B}_0 = 0$; then $\mathfrak{A}_0$ is unitary, by (9). Consequently this mapping has the particular form

$$\mathfrak{W}_0 = \mathfrak{U}'\mathfrak{Z}_0\mathfrak{U},$$

where $\mathfrak{U}$ denotes an arbitrary unitary matrix, i. e., a matrix satisfying $\mathfrak{U}'\bar{\mathfrak{U}} = \mathfrak{E}$. Since (5) maps 0 into $i\mathfrak{E}$, the formula

$$\frac{\mathfrak{W} - i\mathfrak{E}}{\mathfrak{W} + i\mathfrak{E}} = \mathfrak{U}' \frac{\mathfrak{Z} - i\mathfrak{E}}{\mathfrak{Z} + i\mathfrak{E}} \mathfrak{U}$$

gives all symplectic transformations having $i\mathfrak{E}$ as a fixed point.

7. By the results of the preceding section, the proof of Theorem 1 is reduced to the proof of the following statement: *Let $\mathfrak{Z}_0 \to \mathfrak{W}_0$ be an analytic mapping of E onto itself with the fixed point 0; then $\mathfrak{W}_0 = \mathfrak{U}'\mathfrak{Z}_0\mathfrak{U}$ with unitary constant $\mathfrak{U}$.*

Let $\mathfrak{Z}_0 = (z_{kl})$ be an arbitrary point of E and denote by r_k $(k = 1, \cdots, n)$ the characteristic roots of the hermitian matrix $\mathfrak{Z}_0\bar{\mathfrak{Z}}_0$; then $r_k \geq 0$ and also $r_k < 1$, by $\mathfrak{E} - \mathfrak{Z}_0\bar{\mathfrak{Z}}_0 > 0$; we may assume $0 \leq r_1 \leq r_2 \leq \cdots \leq r_n < 1$. The matrix $\mathfrak{Z} = t\mathfrak{Z}_0$ is again a point of E, if the complex scalar factor t satisfies the condition $r_n t\bar{t} < 1$. Let $\mathfrak{W} = \mathfrak{W}(t)$ be the image of $\mathfrak{Z}$ under the analytic mapping $\mathfrak{Z}_0 \to \mathfrak{W}_0$. The elements of the matrix $\mathfrak{W}$ are analytic functions of the single complex variable t, for given $\mathfrak{Z}_0$; they are regular in the circle $r_n t\bar{t} < 1$ and a fortiori in the unit-circle $t\bar{t} \leq 1$. Consequently

$$\mathfrak{W} = \sum_{k=1}^{\infty} t^k \mathfrak{W}_k \qquad (r_n t\bar{t} > 1), \tag{11}$$

where the coefficients $\mathfrak{W}_k$ are matrices depending only upon $\mathfrak{Z}_0$. On the other hand, $\mathfrak{W}$ may be expressed as a power series in the variables tz_{kl}, converging for sufficiently small values of $t\bar{t}$. Since this expansion is unique, the matrix $t^k\mathfrak{W}_k$ is exactly the aggregate of the terms of order k in that power series. This proves, in particular, that the power series representation $\mathfrak{W}_0 = \sum_{k=1}^{\infty} \mathfrak{W}_k$ converges everywhere in E, if we do not split up the polynomials $\mathfrak{W}_k$ into their single terms.

Since $\mathfrak{Z}_0 \to \mathfrak{W}_0$ maps E onto itself, we have $\mathfrak{E} - \mathfrak{W}\bar{\mathfrak{W}} > 0$ for $t\bar{t} = 1$. Integrating over that circle we obtain

$$\frac{1}{2\pi i} \int_{t\bar{t}=1} (\mathfrak{E} - \mathfrak{W}\bar{\mathfrak{W}}) \frac{dt}{t} > 0,$$

whence by (11)

$$\mathfrak{E} - \sum_{k=1}^{\infty} \mathfrak{W}_k \bar{\mathfrak{W}}_k > 0 \tag{12}$$

and in particular

$$\mathfrak{E} - \mathfrak{W}_1 \bar{\mathfrak{W}}_1 > 0. \tag{13}$$

The $\frac{1}{2}n(n+1)$ elements w_{kl} $(1 \leq k \leq l \leq n)$ of $\mathfrak{W}_1 = (w_{kl})$ are linear functions of the independent variables z_{kl} $(1 \leq k \leq l \leq n)$; let D be their determinant. Since $\mathfrak{W}_1$ is the linear part of the power series for $\mathfrak{W}_0$, the functional determinant of the $\frac{1}{2}n(n+1)$ independent elements of $\mathfrak{W}_0$ with respect to the variables z_{kl} is also D, at the point $\mathfrak{Z}_0 = 0$. If we interchange $\mathfrak{W}_0$ and $\mathfrak{Z}_0$, the determinant D is replaced by D^{-1}. In order to prove Theorem 1, we may therefore assume $D\bar{D} \geq 1$.

Consider now the linear mapping $\mathfrak{Z}_0 \to \mathfrak{W}_1$, with the determinant D. By (13), the domain E is mapped onto a domain E_1 contained in E. Let $v(E)$ and $v(E_1)$ be the euclidean volumes of E and E_1, the real and imaginary parts of the z_{kl} being rectangular cartesian coördinates. Then $v(E_1) = D\bar{D}v(E) \geq v(E)$, whence $D\bar{D} = 1$, $E_1 = E$, and the boundary of E is mapped onto itself. We take $\mathfrak{Z}_0 = \mathfrak{U}'\mathfrak{P}\mathfrak{U}$ with unitary $\mathfrak{U}$ and $\mathfrak{P} = [p_1, \cdots, p_n]$; obviously $\mathfrak{Z}_0$ is a boundary point of E, if $-1 \leq p_k \leq 1$ $(k = 1, \cdots, n)$ and at least one $p_k = \pm 1$. On the other hand, the determinant $|\mathfrak{E} - \mathfrak{W}_1\bar{\mathfrak{W}}_1|$ is a polynomial in $p_1, \cdots, p_n$, of total degree $2n$. Since $|\mathfrak{E} - \mathfrak{W}_1\bar{\mathfrak{W}}_1|$ vanishes on the boundary of E, this polynomial is divisible by $\prod_{k=1}^{n}(1 - p_k^2)$, of total degree $2n$. Moreover the constant terms in both polynomials have the value 1; hence

$$|\mathfrak{E} - \mathfrak{W}_1\bar{\mathfrak{W}}_1| = |\mathfrak{E} - \mathfrak{Z}_0\bar{\mathfrak{Z}}_0| \tag{14}$$

for $\mathfrak{Z}_0 = \mathfrak{U}'\mathfrak{P}\mathfrak{U}$, where $\mathfrak{U}$ is an arbitrary unitary matrix and $\mathfrak{P}$ an arbitrary real diagonal matrix. We use now the following lemma, the proof of which will be given in Section **9**.

LEMMA 1. *Let $\mathfrak{Z}$ be a complex symmetric matrix and $\mathfrak{P}$ the diagonal matrix $[q_1^{\frac{1}{2}}, \cdots, q_n^{\frac{1}{2}}]$, where $q_1, \cdots, q_n$ denote the characteristic roots of $\mathfrak{Z}\bar{\mathfrak{Z}}$. There exists a unitary matrix $\mathfrak{U}$ such that $\mathfrak{Z} = \mathfrak{U}'\mathfrak{P}\mathfrak{U}$.*

On account of this lemma, the relationship (14) holds also identically in $\mathfrak{Z}_0 = (z_{kl})$. Since $\mathfrak{W}_1$ is linear in all z_{kl}, we obtain

$$|\lambda\mathfrak{E} - \mathfrak{W}_1\bar{\mathfrak{W}}_1| = |\lambda\mathfrak{E} - \mathfrak{Z}_0\bar{\mathfrak{Z}}_0|$$

identically in λ. This proves that $\mathfrak{Z}_0\bar{\mathfrak{Z}}_0$ and $\mathfrak{W}_1\bar{\mathfrak{W}}_1$ have the same characteristic roots. Applying again Lemma 1, we find

(15) $$\mathfrak{W}_1 = \mathfrak{U}'\mathfrak{Z}_0\mathfrak{U}$$

with unitary $\mathfrak{U}$.

By (12), the inequality $\mathfrak{E} - \mathfrak{W}_1\bar{\mathfrak{W}}_1 - \mathfrak{W}_k\bar{\mathfrak{W}}_k > 0$ holds for $k = 2, 3, \cdots$ and every $\mathfrak{Z}_0$ in E. Choose, in particular, $\mathfrak{Z}_0 = u \exp i\mathfrak{S}$ with real symmetric $\mathfrak{S}$ and $0 < u < 1$. Then, by (15),

$$(1 - u^2)\mathfrak{E} - \mathfrak{W}_k\bar{\mathfrak{W}}_k > 0 \qquad (k = 2, 3, \cdots);$$

hence $\mathfrak{W}_k$ tends to 0, if u tends to 1, and $\mathfrak{W}_k = 0$ for $\mathfrak{Z}_0 = \exp i\mathfrak{S}$ and arbitrary real symmetric $\mathfrak{S}$. But $\mathfrak{W}_k$ is analytic and consequently $\mathfrak{W}_k = 0$ also for $\mathfrak{Z}_0 = \exp i\mathfrak{S}$ with complex symmetric $\mathfrak{S}$. This proves that $\mathfrak{W}_k$ vanishes identically.

8. In order to complete the proof of Theorem 1, we have to prove that the unitary matrix $\mathfrak{U}$ in (15) can be chosen as a constant matrix. Let

(16) $$\mathfrak{W}_1 = \mathfrak{W}_1(\mathfrak{Z}_0) = \sum_{k \le l} z_{kl}\mathfrak{A}_{kl}$$

with constant matrices $\mathfrak{A}_{kl}$ and define

(17) $$\mathfrak{W}_1^* = \mathfrak{W}_1^*(\mathfrak{Z}_0) = \sum_{k \le l} z_{kl}\bar{\mathfrak{A}}_{kl};$$

whence $\bar{\mathfrak{W}}_1 = \mathfrak{W}_1^*(\bar{\mathfrak{Z}}_0)$. Now $\mathfrak{W}_1\bar{\mathfrak{W}}_1 = \mathfrak{E}$ for $\mathfrak{Z}_0\bar{\mathfrak{Z}}_0 = \mathfrak{E}$ and consequently

(18) $$\mathfrak{W}_1(\mathfrak{Z}_0)\mathfrak{W}_1^*(\mathfrak{Z}_0^{-1}) = \mathfrak{E}$$

for $\mathfrak{Z}_0 = \exp i\mathfrak{S}$ with arbitrary real symmetric $\mathfrak{S}$. Since $\mathfrak{W}_1$ and $\mathfrak{W}_1^*$ are analytic, we infer again that (18) is an identity in $\mathfrak{Z}_0$.

Putting $z_{kk} = z_k \neq 0$ $(k = 1, \cdots, n)$, $\mathfrak{Z}_1 = [z_1, \cdots, z_n]$, $\mathfrak{Z}_2 = \mathfrak{Z}_0 - \mathfrak{Z}_1$ and using the Taylor series in the neighborhood of $\mathfrak{Z}_2 = 0$, we find

$$\mathfrak{Z}_0^{-1} = \mathfrak{Z}_1^{-1}(\mathfrak{E} + \mathfrak{Z}_2\mathfrak{Z}_1^{-1})^{-1} = \mathfrak{Z}_1^{-1} - \mathfrak{Z}_1^{-1}\mathfrak{Z}_2\mathfrak{Z}_1^{-1} + \cdots$$
$$(\mathfrak{W}_1(\mathfrak{Z}_1) + \mathfrak{W}_1(\mathfrak{Z}_2))(\mathfrak{W}_1^*(\mathfrak{Z}_1^{-1}) - \mathfrak{W}_1^*(\mathfrak{Z}_1^{-1}\mathfrak{Z}_2\mathfrak{Z}_1^{-1}) + \cdots) = \mathfrak{E},$$

hence in particular

(19) $$\mathfrak{W}_1(\mathfrak{Z}_1)\mathfrak{W}_1^*(\mathfrak{Z}_1^{-1}) = \mathfrak{E},$$

(20) $$\mathfrak{W}_1(\mathfrak{Z}_2)\mathfrak{W}_1^*(\mathfrak{Z}_1^{-1}) = \mathfrak{W}_1(\mathfrak{Z}_1)\mathfrak{W}_1^*(\mathfrak{Z}_1^{-1}\mathfrak{Z}_2\mathfrak{Z}_1^{-1}).$$

It follows from (16), (17) and (19) that

$$\sum_{k,l=1}^{n} z_k z_l^{-1}\mathfrak{A}_k\bar{\mathfrak{A}}_l = \mathfrak{E}$$

with $\mathfrak{A}_k = \mathfrak{A}_{kk}$ $(k = 1, \cdots, n)$, whence

(21) $$\mathfrak{A}_k\bar{\mathfrak{A}}_l = 0 \qquad (k \neq l).$$

By (15), the matrix $\mathfrak{A}_k\bar{\mathfrak{A}}_k$ has the characteristic roots $1, 0, \cdots, 0$. Without loss of generality, we may replace $\mathfrak{W}_1$ by $\mathfrak{U}_1'\mathfrak{W}_1\mathfrak{U}_1$, for any constant unitary matrix $\mathfrak{U}_1$. On account of Lemma 1, we may therefore assume that $\mathfrak{A}_1 = [1, 0, \cdots, 0]$. Then, by (21), the matrices $\mathfrak{A}_2, \cdots, \mathfrak{A}_n$ have the form

$$\mathfrak{A}_k = \begin{pmatrix} 0 & 0 \\ 0 & \mathfrak{B}_k^{(n-1)} \end{pmatrix} \qquad (k = 2, \cdots, n).$$

By induction, applying again Lemma 1 and (21), we may assume

$$\mathfrak{A}_k = [e_{k1}, e_{k2}, \cdots, e_{kn}] \qquad (k = 1, \cdots, n)$$

with $e_{kl} = 0 (k \neq l)$ and $e_{kk} = 1$; hence

$$\mathfrak{W}_1(\mathfrak{Z}_1) = \mathfrak{Z}_1. \tag{22}$$

It follows from (20) and (22) that

$$\mathfrak{W}_1(\mathfrak{Z}_2) = \mathfrak{Z}_1\mathfrak{W}^*_1(\mathfrak{Z}_1^{-1}\mathfrak{Z}_2\mathfrak{Z}_1^{-1})\mathfrak{Z}_1,$$

whence

$$z_k z_l \mathfrak{A}_{kl} = \mathfrak{Z}_1\bar{\mathfrak{A}}_{kl}\mathfrak{Z}_1 \qquad (k \neq l).$$

Consequently

$$\mathfrak{W}_1 = (a_{kl}z_{kl})$$

with real $a_{kl} = a_{lk}$ and $a_{kk} = 1$. Since the matrix $[\pm 1, \pm 1, \cdots, \pm 1]$ is unitary, we may, moreover, assume $a_{1l} \geq 0$ $(l = 2, \cdots, n)$.

The expression $|\mathfrak{W}_1|\,|\mathfrak{Z}_0|^{-1}$ is a rational function of the z_{kl} and has, by (15), the constant absolute value 1; hence it is identically constant. On the other hand, both determinants $|\mathfrak{W}_1|$ and $|\mathfrak{Z}_0|$ contain the term $z_1 z_2 \cdots z_n = z$ with the same coefficient 1. This proves $|\mathfrak{W}_1| = |\mathfrak{Z}_0|$. The term $(z_1 z_l)^{-1} z z_{1l}^2$ $(l = 2, \cdots, n)$ has in $|\mathfrak{W}_1|$ the coefficient $-a_{1l}^2$ and in $|\mathfrak{Z}_0|$ the coefficient -1, hence $a_{1l} = 1$. The term $(z_1 z_k z_l)^{-1} z z_{1k} z_{1l} z_{kl}$ $(1 < k < l)$ has then in $|\mathfrak{W}_1|$ the coefficient $2a_{kl}$ and in $|\mathfrak{Z}_0|$ the coefficient 2, hence $a_{kl} = 1$ and $\mathfrak{W}_1 = \mathfrak{Z}_0$.

9. It remains to prove Lemma 1. There exists a unitary matrix $\mathfrak{U}_1$ such that

$$\mathfrak{Z}\bar{\mathfrak{Z}} = \mathfrak{P}^2\{\mathfrak{U}_1\}.$$

Then the matrix $\mathfrak{Z}[\mathfrak{U}_1^{-1}] = \mathfrak{F}$ is symmetric and $\mathfrak{F}\bar{\mathfrak{F}} = \mathfrak{P}^2$. Let $\mathfrak{F}_1$ and $\mathfrak{F}_2$ be the real and imaginary parts of $\mathfrak{F} = \mathfrak{F}_1 + i\mathfrak{F}_2$. Since $\mathfrak{P}$ is real, we obtain $\mathfrak{F}_1\mathfrak{F}_2 = \mathfrak{F}_2\mathfrak{F}_1$; consequently the two real symmetric matrices $\mathfrak{F}_1$ and $\mathfrak{F}_2$ are permutable. This proves the existence of a real orthogonal matrix $\mathfrak{O}$ such that $\mathfrak{F}_1[\mathfrak{O}]$ and $\mathfrak{F}_2[\mathfrak{O}]$ are both diagonal matrices. Hence $\mathfrak{F}[\mathfrak{O}] = \mathfrak{R}$ is also a diago-

nal matrix $[r_1, \cdots, r_n]$ and $\mathfrak{R}\bar{\mathfrak{R}} = \mathfrak{P}^2[\mathfrak{D}]$. The numbers $r_k\bar{r}_k$ $(k = 1, \cdots, n)$ are therefore a permutation of $q_1, \cdots, q_n$, and we may obviously assume $r_k\bar{r}_k = q_k$.

The diagonal matrix $\mathfrak{U}_2 = [s_1, \cdots, s_n]$ with $s_k = r_k^{1/2} q_k^{-(1/4)}$ $(k = 1, \cdots, n)$ is unitary and $\mathfrak{P}[\mathfrak{U}_2] = \mathfrak{R}$. Defining $\mathfrak{U} = \mathfrak{U}_2\mathfrak{D}'\mathfrak{U}_1$, we have $\mathfrak{P}[\mathfrak{U}] = \mathfrak{R}[\mathfrak{D}'\mathfrak{U}_1]$ $= \mathfrak{F}[\mathfrak{U}_1] = \mathfrak{Z}$; q. e. d.

10. Consider the symplectic mappings in the case $n = 1$, i. e.,

$$w = \frac{az + b}{cz + d} \tag{23}$$

with real a, b, c, d and $ad - bc = 1$. It is well-known that there exists a transformation (23) mapping two given points z, z_1 of the upper half-plane into two other given points w, w_1 of the upper half-plane, if and only if $R(z, z_1) = R(w, w_1)$, where $R(z, z_1)$ denotes the cross-ratio $\frac{z - z_1}{z - \bar{z}_1}\frac{\bar{z} - \bar{z}_1}{\bar{z} - z_1}$. Theorem 2 is the generalization to the case of an arbitrary n.

Let $\mathfrak{Z}, \mathfrak{Z}_1$ be two points of H and $\mathfrak{W}, \mathfrak{W}_1$ their images under the symplectic mapping $\mathfrak{W} = (\mathfrak{A}\mathfrak{Z} + \mathfrak{B})(\mathfrak{C}\mathfrak{Z} + \mathfrak{D})^{-1}$ with the matrix $\mathfrak{M}$. We have

$$\begin{aligned} \mathfrak{Z}_1 - \mathfrak{Z} &= (\mathfrak{Z}_1\mathfrak{E})\mathfrak{J}\begin{pmatrix}\mathfrak{Z}\\ \mathfrak{E}\end{pmatrix} = (\mathfrak{Z}_1\mathfrak{E})\mathfrak{M}'\mathfrak{J}\mathfrak{M}\begin{pmatrix}\mathfrak{Z}\\ \mathfrak{E}\end{pmatrix} \\ &= (\mathfrak{C}\mathfrak{Z}_1 + \mathfrak{D})'(\mathfrak{W}_1 - \mathfrak{W})(\mathfrak{C}\mathfrak{Z} + \mathfrak{D}), \end{aligned} \tag{24}$$

$$\bar{\mathfrak{Z}}_1 - \mathfrak{Z} = (\mathfrak{C}\bar{\mathfrak{Z}}_1 + \mathfrak{D})'(\bar{\mathfrak{W}}_1 - \mathfrak{W})(\mathfrak{C}\mathfrak{Z} + \mathfrak{D}). \tag{25}$$

Now $(\mathfrak{Z} - \bar{\mathfrak{Z}}_1)^{-1}$ exists, since $\mathfrak{Z} - \bar{\mathfrak{Z}}_1$ is a point of H; consequently

$$\begin{aligned} &(\mathfrak{Z} - \mathfrak{Z}_1)(\mathfrak{Z} - \bar{\mathfrak{Z}}_1)^{-1} \\ &\quad = (\mathfrak{Z}_1\mathfrak{C}' + \mathfrak{D}')(\mathfrak{W} - \mathfrak{W}_1)(\mathfrak{W} - \bar{\mathfrak{W}}_1)^{-1}(\bar{\mathfrak{Z}}_1\mathfrak{C}' + \mathfrak{D}')^{-1}. \end{aligned}$$

Introducing the cross-ratio

$$\mathfrak{R} = \mathfrak{R}(\mathfrak{Z}, \mathfrak{Z}_1) = (\mathfrak{Z} - \mathfrak{Z}_1)(\mathfrak{Z} - \bar{\mathfrak{Z}}_1)^{-1}(\bar{\mathfrak{Z}} - \bar{\mathfrak{Z}}_1)(\bar{\mathfrak{Z}} - \mathfrak{Z}_1)^{-1}$$

and putting $\mathfrak{R}^* = \mathfrak{R}(\mathfrak{W}, \mathfrak{W}_1)$, $\mathfrak{Q} = (\mathfrak{C}\mathfrak{Z}_1 + \mathfrak{D})'$, we find

$$\mathfrak{R} = \mathfrak{Q}\mathfrak{R}^*\mathfrak{Q}^{-1}. \tag{26}$$

Hence the matrices $\mathfrak{R}$ and $\mathfrak{R}^*$ have the same characteristic roots.

Choose in particular $\mathfrak{Z}_1 = i\mathfrak{E}$, $\mathfrak{Z} = i\mathfrak{T}$ with $\mathfrak{T} = [t_1, \cdots, t_n]$ and $1 \leq t_1 \leq t_2 \leq \cdots \leq t_n$. Then $\mathfrak{R} = [r_1, \cdots, r_n]$ with

$$r_k = \left(\frac{t_k - 1}{t_k + 1}\right)^2 \qquad (k = 1, \cdots n),$$

whence $0 \leq r_1 \leq r_2 \leq \cdots \leq r_n < 1$ and

$$t_k = \frac{1 + r_k^{\frac{1}{2}}}{1 - r_k^{\frac{1}{2}}}. \tag{27}$$

In this case, the diagonal matrix $\mathfrak{T}$ is uniquely determined by the characteristic roots $r_1, \cdots, r_n$ of $\mathfrak{R}$. In order to complete the proof of Theorem 2, we have only to prove

LEMMA 2. *Let $\mathfrak{Z}, \mathfrak{Z}_1$ be two arbitrary points of H. There exists a symplectic transformation mapping $\mathfrak{Z}, \mathfrak{Z}_1$ into $i\mathfrak{T}, i\mathfrak{E}$ with $\mathfrak{T} = [t_1, \cdots, t_n]$ and $1 \leq t_1 \leq t_2 \leq \cdots \leq t_n$.*

Since Ω is transitive, we may already assume $\mathfrak{Z}_1 = i\mathfrak{E}$. If $\mathfrak{T} = [t_1, \cdots, t_n]$ with $1 \leq t_1 \leq t_2 \leq \cdots \leq t_n$, then $(\mathfrak{T} - \mathfrak{E})(\mathfrak{T} + \mathfrak{E})^{-1} = \mathfrak{P} = [p_1, \cdots, p_n]$ with $p_k = (t_k - 1)(t_k + 1)^{-1}$ $(k = 1, \cdots, n)$, $0 \leq p_1 \leq p_2 \leq \cdots \leq p_n < 1$, and vice versa. By (5), (6) and the results of Section **6**, we have only to prove that there exists for every point $\mathfrak{Z}$ of E a unitary matrix $\mathfrak{U}_1$ satisfying $\mathfrak{U}_1' \mathfrak{Z} \mathfrak{U}_1 = \mathfrak{P} = [p_1, \cdots, p_n]$ with $0 \leq p_1 \leq p_2 \leq \cdots \leq p_n < 1$. This follows from Lemma 1: We choose $\mathfrak{U}_1 = \mathfrak{U}^{-1}$ and $p_k = q_k^{\frac{1}{2}}$; since $q_1, \cdots, q_n$ are the characteristic roots of the hermitian matrix $\mathfrak{Z}\bar{\mathfrak{Z}}$ and $\mathfrak{E} - \mathfrak{Z}\bar{\mathfrak{Z}} > 0$, we may assume $0 \leq q_1 \leq q_2 \leq \cdots \leq q_n < 1$, and $p_1, \cdots, p_n$ have the required property.

On account of the symplectic invariance of the characteristic roots $r_1, \cdots, r_n$ of $\mathfrak{R}(\mathfrak{Z}, \mathfrak{Z}_1)$, the diagonal elements t_k of $\mathfrak{T}$ are given by (27). This proves that those characteristic roots are always real numbers of the interval $0 \leq r < 1$.

III. THE SYMPLECTIC METRIC.

11. We consider the cross-ratio

$$\mathfrak{R} = (\mathfrak{Z} - \mathfrak{Z}_1)(\mathfrak{Z} - \bar{\mathfrak{Z}}_1)^{-1}(\bar{\mathfrak{Z}} - \bar{\mathfrak{Z}}_1)(\bar{\mathfrak{Z}} - \mathfrak{Z}_1)^{-1}$$

as a function of $\mathfrak{Z}_1$, for any given $\mathfrak{Z}$ in H. Since the two factors $\mathfrak{Z} - \mathfrak{Z}_1$ and $\bar{\mathfrak{Z}} - \bar{\mathfrak{Z}}_1$ vanish for $\mathfrak{Z}_1 = \mathfrak{Z}$, the second differential of $\mathfrak{R}$, at the point $\mathfrak{Z}_1 = \mathfrak{Z}$, has the value

$$d^2\mathfrak{R} = 2d\mathfrak{Z}(\mathfrak{Z} - \bar{\mathfrak{Z}})^{-1} d\bar{\mathfrak{Z}}(\bar{\mathfrak{Z}} - \mathfrak{Z})^{-1} = \tfrac{1}{2} d\mathfrak{Z}\mathfrak{Y}^{-1} d\bar{\mathfrak{Z}}\mathfrak{Y}^{-1},$$

where $\mathfrak{Y}$ denotes the imaginary part of $\mathfrak{Z} = \mathfrak{X} + i\mathfrak{Y}$. On the other hand, by (26), the trace $\sigma(\mathfrak{R})$ of $\mathfrak{R} = \mathfrak{R}(\mathfrak{Z}, \mathfrak{Z}_1)$ is invariant under any cogredient

symplectic transformation of the points $\mathfrak{Z}, \mathfrak{Z}_1$. Moreover $d^2\sigma(\mathfrak{R}) = \sigma(d^2\mathfrak{R})$, and consequently the hermitian differential form

$$ds^2 = \sigma(\mathfrak{Y}^{-1}d\mathfrak{Z}\mathfrak{Y}^{-1}d\bar{\mathfrak{Z}}) \tag{28}$$

is invariant under Ω. Introducing $\mathfrak{X} = (x_{kl})$ and $\mathfrak{Y} = (y_{kl})$, we obtain

$$ds^2 = \sigma(\mathfrak{Y}^{-1}d\mathfrak{X}\mathfrak{Y}^{-1}d\mathfrak{X} + \mathfrak{Y}^{-1}d\mathfrak{Y}\mathfrak{Y}^{-1}d\mathfrak{Y}) \tag{29}$$

and in particular, for $\mathfrak{Z} = i\mathfrak{E}$,

$$ds^2 = \sum_{k=1}^{n}(dx_{kk}^2 + dy_{kk}^2) + 2\sum_{k<l}(dx_{kl}^2 + dy_{kl}^2).$$

Since Ω is transitive in H, the quadratic differential form ds^2 is obviously positive definite everywhere in H.

Let us determine the most general quadratic differential invariant Q of the symplectic group. On account of the transitivity of Ω, we have only to find Q at the point $\mathfrak{Z} = i\mathfrak{E}$ of H or, if we use the variable $\mathfrak{Z}_0 = (\mathfrak{Z} - i\mathfrak{E})(\mathfrak{Z} + i\mathfrak{E})^{-1}$ already defined in (6), at the point $\mathfrak{Z}_0 = 0$ of E. Then Q becomes a quadratic form of the elements of $\mathfrak{S} = d\mathfrak{Z}_0$ and $\bar{\mathfrak{S}} = d\bar{\mathfrak{Z}}_0$ which is, by the result of Section **6**, invariant under all transformations $\mathfrak{S} \to \mathfrak{U}'\mathfrak{S}\mathfrak{U}$ with unitary $\mathfrak{U}$. By Lemma 1, there exists for any complex symmetric $\mathfrak{S}$ a unitary matrix $\mathfrak{U}$, such that $\mathfrak{U}'\mathfrak{S}\mathfrak{U} = \mathfrak{P} = [p_1, \cdots, p_n]$ with real p_k $(k = 1, \cdots, n)$, where the p_k^2 are the characteristic roots of $\mathfrak{S}\bar{\mathfrak{S}}$. This proves that Q is a quadratic function of $p_1, \cdots, p_n$ alone. Let $k_1, \cdots, k_n$ be a permutation of the numbers $1, \cdots, n$ and ϵ_l $(l = 1, \cdots, n)$ a fourth root of unity; then the matrix $\mathfrak{U}_0$ of the substitution $s_{k_l} \to \epsilon_l s_l$ $(l = 1, \cdots, n)$ is unitary and $\mathfrak{U}_0'\mathfrak{P}\mathfrak{U}_0 = [q_1, \cdots, q_n]$ with $q_l = \epsilon_l^2 p_{k_l} = \pm p_{k_l}$ $(l = 1, \cdots, n)$. Hence Q is a symmetric polynomial in $p_1^2, \cdots, p_n^2$,

$$Q = \lambda\sum_{k=1}^{n} p_k^2 = \lambda\sigma(\mathfrak{S}\bar{\mathfrak{S}})$$

with constant λ. Consequently any quadratic differential invariant of the symplectic group is a constant multiple of ds^2.

12. We are now interested in the properties of the geodesics for the symplectic metric (28). In order to find the shortest arc connecting two arbitrarily given points $\mathfrak{Z}_1$ and $\mathfrak{Z}_2$ of H, we have, by Lemma 2, only to investigate the special case $\mathfrak{Z}_1 = i\mathfrak{E}$, $\mathfrak{Z}_2 = i\mathfrak{T} = i[t_1, \cdots, t_n]$, $1 \leq t_1 \leq t_2 \leq \cdots \leq t_n$; moreover, we may obviously assume $\mathfrak{Z}_1 \neq \mathfrak{Z}_2$, i. e., $t_n > 1$. Let now $\mathfrak{Z} = \mathfrak{Z}(u)$ be any curve connecting these two points in H and having a piecewise continuous tangent, $\mathfrak{Z}(0) = i\mathfrak{E}$, $\mathfrak{Z}(1) = i\mathfrak{T}$. We may put

$$\mathfrak{Z} = \mathfrak{X} + i\mathfrak{Q}[\mathfrak{D}],$$

where $\mathfrak{Q} = [q_1, \cdots, q_n]$ with $q_k > 0$ $(k = 1, \cdots, n)$ and $\mathfrak{D}$ denotes a real orthogonal matrix; moreover $\mathfrak{X}$, $\mathfrak{Q}$, $\mathfrak{D}$ have again piecewise continuous derivatives; $\mathfrak{Q} = \mathfrak{E}$, $\mathfrak{D} = \mathfrak{E}$, $\mathfrak{X} = 0$ for $u = 0$; $\mathfrak{Q} = \mathfrak{T}$, $\mathfrak{D} = \mathfrak{E}$, $\mathfrak{X} = 0$ for $u = 1$. This arc has the length

$$s = \int_0^1 \sigma^{\frac{1}{2}}(\mathfrak{Y}^{-1}\dot{\mathfrak{Z}}\mathfrak{Y}^{-1}\dot{\bar{\mathfrak{Z}}})\, du,$$

where the dot denotes differentiation with respect to u.

By (29), we have $s \geq s_1$, where s_1 denotes the length of the curve

$$\mathfrak{Z} = i\mathfrak{Q}[\mathfrak{D}] \tag{30}$$

also connecting $\mathfrak{Z}_1$ and $\mathfrak{Z}_2$, and $s > s_1$, if both curves do not coincide.

We use the abbreviation $\dot{\mathfrak{D}}\mathfrak{D}' = \mathfrak{F} = (f_{kl})$. Since $\mathfrak{D}\mathfrak{D}' = \mathfrak{E}$, we have $\dot{\mathfrak{D}}\mathfrak{D}' = -\mathfrak{D}\dot{\mathfrak{D}}'$, and consequently $\mathfrak{F}$ is skew-symmetric. Differentiating the equation $\mathfrak{D}\mathfrak{Y}\mathfrak{D}' = \mathfrak{Q}$, we obtain

$$\mathfrak{D}\dot{\mathfrak{Y}}\mathfrak{D}' = \dot{\mathfrak{Q}} - \mathfrak{F}\mathfrak{Q} + \mathfrak{Q}\mathfrak{F}$$

$$\begin{aligned}
\mathfrak{D}\mathfrak{Y}^{-1}\dot{\mathfrak{Y}}\mathfrak{Y}^{-1}\dot{\mathfrak{Y}}\mathfrak{D}' &= \mathfrak{Q}^{-1}(\dot{\mathfrak{Q}} - \mathfrak{F}\mathfrak{Q} + \mathfrak{Q}\mathfrak{F})\mathfrak{Q}^{-1}(\dot{\mathfrak{Q}} - \mathfrak{F}\mathfrak{Q} + \mathfrak{Q}\mathfrak{F}) \\
&= \mathfrak{F}^2 + \mathfrak{Q}^{-1}\mathfrak{F}^2\mathfrak{Q} + \mathfrak{Q}^{-1}\dot{\mathfrak{Q}}\mathfrak{Q}^{-1}\dot{\mathfrak{Q}} - \mathfrak{Q}^{-1}\mathfrak{F}\mathfrak{Q}\mathfrak{F} - \mathfrak{F}\mathfrak{Q}^{-1}\mathfrak{F}\mathfrak{Q} \\
&\quad - \mathfrak{Q}^{-1}\dot{\mathfrak{Q}}\mathfrak{Q}^{-1}\mathfrak{F}\mathfrak{Q} - \mathfrak{Q}^{-1}\mathfrak{F}\dot{\mathfrak{Q}} + \mathfrak{Q}^{-1}\dot{\mathfrak{Q}}\mathfrak{F} + \mathfrak{F}\mathfrak{Q}^{-1}\dot{\mathfrak{Q}} \\
\sigma(\mathfrak{Y}^{-1}\dot{\mathfrak{Y}}\mathfrak{Y}^{-1}\dot{\mathfrak{Y}}) &= 2\sigma(\mathfrak{F}^2) - 2\sigma(\mathfrak{F}\mathfrak{Q}\mathfrak{F}\mathfrak{Q}^{-1}) + \sigma(\mathfrak{Q}^{-1}\dot{\mathfrak{Q}}\mathfrak{Q}^{-1}\dot{\mathfrak{Q}}) \\
&= \sum_{k,l=1}^{n} f_{kl}{}^2 \frac{(q_k - q_l)^2}{q_k q_l} + \sum_{k=1}^{n}\left(\frac{\dot{q}_k}{q_k}\right)^2. \qquad (31)
\end{aligned}$$

On the other hand, the formula

$$\sum_{k=1}^{n} Q_k{}^2 = (c_1 Q_1 + \cdots + c_n Q_n)^2 + \tfrac{1}{2}\sum_{k,l=1}^{n}(c_k Q_l - c_l Q_k)^2 \tag{32}$$

holds for $c_1{}^2 + c_2{}^2 + \cdots + c_n{}^2 = 1$ and in particular with

$$Q_k = \frac{\dot{q}_k}{q_k}, \qquad c_k = \rho^{-1}\log t_k \qquad (k = 1, \cdots, n),$$

where

$$\rho = (\sum_{k=1}^{n} \log^2 t_k)^{\frac{1}{2}} > 0. \tag{33}$$

By (31) and (32),

$$s_1 \geq \int_0^1 (c_1 Q_1 + \cdots + c_n Q_n)\, du = \sum_{k=1}^{n} c_k \log t_k = \rho,$$

with the sign of inequality, if not all the three conditions

$$f_{kl}(q_k - q_l) = 0 \qquad (k, l = 1, \cdots, n), \tag{34}$$

$$c_k Q_l - c_l Q_k = 0 \qquad (k, l = 1, \cdots, n), \tag{35}$$

$$\sum_{k=1}^{n} c_k Q_k \geq 0 \tag{36}$$

are fulfilled. By (35),

$$\log t_n \log q_k = \log t_k \log q_n \qquad (k = 1, \cdots, n),$$

whence

$$q_k = t_k^{\gamma} \qquad (k = 1, \cdots, n) \tag{37}$$

with $\gamma = \gamma(u)$, $\gamma(0) = 0$, $\gamma(1) = 1$. By (36), the function $\gamma(u)$ is monotone. By (34) and (37),

$$f_{kl}(t_k - t_l) = 0 \qquad (k, l = 1, \cdots, n)$$

$$\mathfrak{D}\dot{\mathfrak{D}}'\mathfrak{T} + \mathfrak{T}\dot{\mathfrak{D}}\mathfrak{D}' = 0$$

$$(\mathfrak{D}'\mathfrak{T}\mathfrak{D})^{\cdot} = \dot{\mathfrak{D}}'\mathfrak{T}\mathfrak{D} + \mathfrak{D}'\mathfrak{T}\dot{\mathfrak{D}} = 0$$

$$\mathfrak{D}'\mathfrak{T}\mathfrak{D} = \mathfrak{T};$$

consequently $\mathfrak{T}\mathfrak{D} = \mathfrak{D}\mathfrak{T}$ and, by (37), also $\mathfrak{Q}\mathfrak{D} = \mathfrak{D}\mathfrak{Q}$. This proves, by (30), that the minimum ρ of s is attained, if and only if $\mathfrak{Z} = i[t_1^{\gamma}, \cdots, t_n^{\gamma}]$, where $\gamma(u)$ is a monotone function with $\gamma(0) = 0$, $\gamma(1) = 1$. We may replace $\gamma(u)$ by u and obtain the curve $\mathfrak{Z} = i[t_1^{u}, \cdots, t_n^{u}]$ as the unique solution. Introducing the length of arc $\tau = \rho u$, we have $\mathfrak{Z} = i[e^{c_1\tau}, \cdots, e^{c_n\tau}]$ with $c_k = \rho^{-1} \log t_k$ $(k = 1, \cdots, n)$ and $c_1^2 + \cdots + c_n^2 = 1$.

13. Let $\mathfrak{Z}$ and $\mathfrak{Z}_1$ be again two arbitrary points of H. By (33) and the results of Section **10**, the symplectic distance $\rho = \rho(\mathfrak{Z}, \mathfrak{Z}_1)$ is given by

$$\rho^2 = \sum_{k=1}^{n} \log^2 \frac{1 + r_k^{\frac{1}{2}}}{1 - r_k^{\frac{1}{2}}}, \tag{38}$$

where $r_1, \cdots, r_n$ denote the characteristic roots of the cross-ratio $\mathfrak{R} = \mathfrak{R}(\mathfrak{Z}, \mathfrak{Z}_1)$. Since

$$\log^2 \frac{1 + r^{\frac{1}{2}}}{1 - r^{\frac{1}{2}}} = 4r \left(\sum_{k=0}^{\infty} \frac{r^k}{2k + 1} \right)^2 \qquad (0 \leq r < 1)$$

and

$$\sum_{k=1}^{n} r_k^{\,l} = \sigma(\mathfrak{R}^l) \qquad (l = 1, 2, \cdots),$$

we may write

$$\rho^2 = \sigma \left(\log^2 \frac{1 + \mathfrak{R}^{\frac{1}{2}}}{1 - \mathfrak{R}^{\frac{1}{2}}} \right)$$

with

$$\log^2 \frac{1+\mathfrak{R}^{\frac{1}{2}}}{1-\mathfrak{R}^{\frac{1}{2}}} = 4\mathfrak{R}\left(\sum_{k=0}^{\infty} \frac{\mathfrak{R}^k}{2k+1}\right)^2.$$

By the results of this and the preceding section, Theorem 3 and Theorem 4 are completely proved.

14. In order to calculate the differential equation of the geodesics and the tensor of curvature, we determine the first variation δs. We consider any curve $\mathfrak{Z} = \mathfrak{Z}(s)$ $(0 \leq s \leq s_0)$, where the parameter s denotes the length of arc. Then $\sigma(\mathfrak{Y}^{-1}\dot{\mathfrak{Z}}\mathfrak{Y}^{-1}\dot{\bar{\mathfrak{Z}}}) = 1$ and

$$\delta s = \tfrac{1}{2}\int_0^{s_0} \delta\sigma(\mathfrak{Y}^{-1}\dot{\mathfrak{Z}}\mathfrak{Y}^{-1}\dot{\bar{\mathfrak{Z}}})\,ds.$$

Using the abbreviations $\dot{\mathfrak{Z}}\mathfrak{Y}^{-1}\dot{\bar{\mathfrak{Z}}} = \mathfrak{V}$, $\mathfrak{Y}^{-1}\dot{\mathfrak{Z}}\mathfrak{Y}^{-1} = \mathfrak{W}$ and denoting by R the real part, we find

$$\delta\sigma(\mathfrak{Y}^{-1}\dot{\mathfrak{Z}}\mathfrak{Y}^{-1}\dot{\bar{\mathfrak{Z}}}) = 2R\sigma(\mathfrak{V}\delta\mathfrak{Y}^{-1} + (\mathfrak{W}\delta\bar{\mathfrak{Z}})^{\cdot} - \dot{\mathfrak{W}}\delta\bar{\mathfrak{Z}})$$

$$\delta\mathfrak{Y}^{-1} = \frac{i}{2}\mathfrak{Y}^{-1}(\delta\mathfrak{Z} - \delta\bar{\mathfrak{Z}})\mathfrak{Y}^{-1}$$

$$\dot{\mathfrak{W}} = \mathfrak{Y}^{-1}\ddot{\mathfrak{Z}}\mathfrak{Y}^{-1} + i\mathfrak{Y}^{-1}\dot{\mathfrak{Z}}\mathfrak{Y}^{-1}\dot{\mathfrak{Z}}\mathfrak{Y}^{-1} - \frac{i}{2}\mathfrak{Y}^{-1}(\mathfrak{V} + \bar{\mathfrak{V}})\mathfrak{Y}^{-1},$$

whence

$$\delta s = -R\int_0^{s_0} \sigma(\mathfrak{Y}^{-1}(\ddot{\mathfrak{Z}} + i\dot{\mathfrak{Z}}\mathfrak{Y}^{-1}\dot{\mathfrak{Z}})\mathfrak{Y}^{-1}\delta\bar{\mathfrak{Z}})\,ds.$$

Consequently

$$\ddot{\mathfrak{Z}} = -i\dot{\mathfrak{Z}}\mathfrak{Y}^{-1}\dot{\mathfrak{Z}} \tag{39}$$

is the differential equation of the geodesic lines.

It is easy to perform the integration without using Section **12**. On account of the symplectic invariance of the geodesics and the transitivity of Ω, it is sufficient to integrate (39) for the initial point $\mathfrak{Z} = i\mathfrak{E}$ and an arbitrary direction $\dot{\mathfrak{Z}} = \mathfrak{R}$ through this point; obviously $\sigma(\mathfrak{R}\bar{\mathfrak{R}}) = 1$. By Section **6**, the mapping

$$\frac{\mathfrak{Z} - i\mathfrak{E}}{\mathfrak{Z} + i\mathfrak{E}} \to \mathfrak{U}'\frac{\mathfrak{Z} - i\mathfrak{E}}{\mathfrak{Z} + i\mathfrak{E}}\mathfrak{U} \tag{40}$$

is symplectic for arbitrary unitary $\mathfrak{U}$. Under this mapping, the direction $\mathfrak{R}$ through $i\mathfrak{E}$ is replaced by $\mathfrak{U}'\mathfrak{R}\mathfrak{U}$. By Lemma 1, we can determine $\mathfrak{U}$ such that $\mathfrak{U}'\mathfrak{R}\mathfrak{U} = i\mathfrak{G} = i[g_1, \cdots, g_n]$ with $0 \leq g_1 \leq \cdots \leq g_n$; since $\sigma(\mathfrak{R}\bar{\mathfrak{R}}) = 1$, we have $g_1^2 + \cdots + g_n^2 = 1$. It is now sufficient to integrate (39) for the

initial conditions $\mathfrak{Z} = i\mathfrak{E}$, $\dot{\mathfrak{Z}} = i\mathfrak{G}$. Obviously the solution is $\mathfrak{Z} = i \exp s\mathfrak{G}$. Consequently the most general geodesic line is

$$(\mathfrak{A}\mathfrak{Z} + \mathfrak{B})(\mathfrak{C}\mathfrak{Z} + \mathfrak{D})^{-1} = i[e^{g_1 s}, \cdots, e^{g_n s}],$$

where the left-hand side is an arbitrary symplectic transformation of $\mathfrak{Z}$ and $g_1, \cdots, g_n$ are arbitrary real numbers with $0 \leq g_1 \leq \cdots \leq g_n$, $g_1^2 + \cdots + g_n^2 = 1$.

15. We shall now establish directly, without using Section **12**, that there exists exactly one geodesic through two arbitrarily given points $\mathfrak{Z}_1$ and $\mathfrak{Z}_2 \neq \mathfrak{Z}_1$ of H. We may again assume $\mathfrak{Z}_1 = i\mathfrak{E}$, $\mathfrak{Z}_2 = i\mathfrak{T} = i[t_1, \cdots, t_n]$, $1 \leq t_1 \leq \cdots \leq t_n$,

$$\rho = (\sum_{k=1}^{n} \log^2 t_k)^{\frac{1}{2}} > 0.$$

Defining $g_k = \rho^{-1} \log t_k$ $(k = 1, \cdots, n)$, we obtain the geodesic line

$$\mathfrak{Z} = i \exp s\mathfrak{G}. \tag{41}$$

Since (40) is the most general symplectic mapping with the fixed point $i\mathfrak{E}$, any geodesic through this point has an equation

$$\frac{\mathfrak{Z} - i\mathfrak{E}}{\mathfrak{Z} + i\mathfrak{E}} = \mathfrak{U}' \frac{\exp s\mathfrak{H} - \mathfrak{E}}{\exp s\mathfrak{H} + \mathfrak{E}} \mathfrak{U} \tag{42}$$

with unitary $\mathfrak{U}$ and $\mathfrak{H} = [h_1, \cdots, h_n]$, $0 \leq h_1 \leq \cdots \leq h_n$, $h_1^2 + \cdots + h_n^2 = 1$. If this curve goes also through the point $\mathfrak{Z} = i\mathfrak{T}$ of (41), we obtain

$$\mathfrak{U}\left(\frac{\mathfrak{E} - \mathfrak{T}}{\mathfrak{E} + \mathfrak{T}}\right)^2 = \left(\frac{\mathfrak{E} - \exp s_0\mathfrak{H}}{\mathfrak{E} + \exp s_0\mathfrak{H}}\right)^2 \mathfrak{U}$$

for a certain $s_0 > 0$, whence $\mathfrak{T} = \exp s_0\mathfrak{H}$, $s_0\mathfrak{H} = \rho\mathfrak{G}$, $s_0 = \rho$, $\mathfrak{H} = \mathfrak{G}$. Moreover $\mathfrak{U}$ and $(\exp s_0\mathfrak{H} - \mathfrak{E})(\exp s_0\mathfrak{H} + \mathfrak{E})^{-1}$ are permutable, hence also $\mathfrak{U}$ and $(\exp s\mathfrak{H} - \mathfrak{E})(\exp s\mathfrak{H} + \mathfrak{E})^{-1}$. Putting $s = \rho$ in (42), we find $\mathfrak{U}'\mathfrak{U} = \mathfrak{E}$, and consequently the geodesics (41) and (42) coincide.

This result proves again, by a general theorem from the calculus of variations, that there exists exactly one shortest arc between two arbitrarily given points of H.

16. By the minimum property of the geodesic arc, the symplectic distance $\rho(\mathfrak{Z}_1, \mathfrak{Z}_2)$ satisfies the triangle inequality

$$\rho(\mathfrak{Z}_1, \mathfrak{Z}_3) \leq \rho(\mathfrak{Z}_1, \mathfrak{Z}_2) + \rho(\mathfrak{Z}_2, \mathfrak{Z}_3)$$

for three arbitrary points $\mathfrak{Z}_1, \mathfrak{Z}_2, \mathfrak{Z}_3$ of H, and the sign of equality is true, if and only if $\mathfrak{Z}_2$ is a point on the uniquely determined geodesic arc between $\mathfrak{Z}_1$

and $\mathfrak{Z}_3$. Obviously $\rho(\mathfrak{Z}, \mathfrak{Z}_1)$ is a continuous function of $\mathfrak{Z}$. If G is any compact point set in H, then $\rho(\mathfrak{Z}, \mathfrak{Z}_1) \leq c$ for all $\mathfrak{Z}$ in G, where c is a positive constant depending only upon $\mathfrak{Z}_1$ and G. Let us prove that also the converse statement is true: If c is an arbitrary positive constant and $\mathfrak{Z}_1$ any given point of H, then the inequality $\rho(\mathfrak{Z}, \mathfrak{Z}_1) \leq c$ defines a compact set G of points $\mathfrak{Z}$ in H. It is sufficient to prove this in the special case $\mathfrak{Z}_1 = i\mathfrak{E}$. By the definition of the cross-ratio,

$$\mathfrak{R}(\mathfrak{Z}, i\mathfrak{E}) = \mathfrak{Z}_0\bar{\mathfrak{Z}}_0,$$

where $\mathfrak{Z}_0 = (\mathfrak{Z} - i\mathfrak{E})(\mathfrak{Z} + i\mathfrak{E})^{-1}$ is the image of $\mathfrak{Z}$ under the transformation (6) mapping H onto E. We infer from (38) that the characteristic roots r_k of the hermitian matrix $\mathfrak{Z}_0\bar{\mathfrak{Z}}_0$ satisfy an inequality

$$0 \leq r_k \leq \vartheta < 1 \qquad (k = 1, \cdots, n)$$

for all $\mathfrak{Z}$ with $\rho(\mathfrak{Z}, \mathfrak{Z}_1) \leq c$, where ϑ depends only upon c. Since an arbitrary symmetric matrix $\mathfrak{Z}_0$ is a point of E, if the characteristic roots of $\mathfrak{Z}_0\bar{\mathfrak{Z}}_0$ are < 1, all limit points of the images $\mathfrak{Z}_0$ of the points $\mathfrak{Z}$ in G belong to E again, and consequently G is compact.

17. Let

$$ds^2 = \sum_{k,l=1}^{m} g_{kl}dx_k dx_l \tag{43}$$

be a Riemann metric in an m-dimensional space and let

$$\ddot{x}_k = -\sum_{p,q=1}^{m} \{pq, k\}\dot{x}_p\dot{x}_q \qquad (k = 1, \cdots, m)$$

be the differential equations of the geodesics, so that $\{pq, k\}$ denotes the Christoffel symbol of the second kind. The Riemann tensor of curvature R is obtained in the following way: Define two covariant differentials $\delta_1 u_k$ and $\delta_2 u_k$ by

$$\delta_r u_k = -\sum_{p,q} \{pq, k\}u_p\delta_r x_q \qquad (r = 1, 2); \tag{44}$$

then

$$R = \sum_{k,l} g_{kl}v_k(\delta_1\delta_2 - \delta_2\delta_1)u_l = \sum_{k,l,p,q} R_{klpq}u_k v_l \delta_1 x_p \delta_2 x_q, \tag{45}$$

where $u_k, v_l, \delta_1 x_p, \delta_2 x_q$ are the components of 4 covariant vectors.

In the case of the symplectic metric (28), the differential equations of the geodesics are given by (39). Instead of (44), we may write now

$$2i\delta_r\mathfrak{U} = \mathfrak{U}\mathfrak{Y}^{-1}\delta_r\mathfrak{Z} + \delta_r\mathfrak{Z}\mathfrak{Y}^{-1}\mathfrak{U} \qquad (r = 1, 2)$$

with complex symmetric $\mathfrak{U} = (u_{kl})$, and we obtain

$$-4\delta_1\delta_2\mathfrak{U} = (\mathfrak{U}\mathfrak{Y}^{-1}\delta_1\bar{\mathfrak{Z}} + \delta_1\mathfrak{Z}\mathfrak{Y}^{-1}\mathfrak{U})\mathfrak{Y}^{-1}\delta_2\mathfrak{Z} + \delta_2\mathfrak{Z}\mathfrak{Y}^{-1}(\mathfrak{U}\mathfrak{Y}^{-1}\delta_1\mathfrak{Z} + \delta_1\bar{\mathfrak{Z}}\mathfrak{Y}^{-1}\mathfrak{U})$$

$$4(\delta_1\delta_2 - \delta_2\delta_1)\mathfrak{U} = \mathfrak{F}\mathfrak{Y}^{-1}\mathfrak{U} + \mathfrak{U}\mathfrak{Y}^{-1}\mathfrak{F}', \tag{46}$$

where

$$\mathfrak{F} = \delta_1\mathfrak{Z}\mathfrak{Y}^{-1}\delta_2\bar{\mathfrak{Z}} - \delta_2\mathfrak{Z}\mathfrak{Y}^{-1}\delta_1\bar{\mathfrak{Z}}. \tag{47}$$

Introducing

$$\mathfrak{G} = \mathfrak{U}\mathfrak{Y}^{-1}\bar{\mathfrak{B}} - \mathfrak{B}\mathfrak{Y}^{-1}\bar{\mathfrak{U}} \tag{48}$$

with complex symmetric $\mathfrak{B} = (v_{kl})$, we find, by (28), (45) and (46),

$$\begin{aligned} 4R &= \sigma(\mathfrak{Y}^{-1}\mathfrak{B}\bar{\mathfrak{F}}\mathfrak{Y}^{-1}\bar{\mathfrak{U}} + \mathfrak{Y}^{-1}\mathfrak{F}\mathfrak{Y}^{-1}\mathfrak{U}\mathfrak{Y}^{-1}\bar{\mathfrak{B}}) \\ &= \sigma(\mathfrak{Y}^{-1}\mathfrak{F}\mathfrak{Y}^{-1}\mathfrak{G}) = -\sigma(\mathfrak{Y}^{-1}\mathfrak{F}\mathfrak{Y}^{-1}\bar{\mathfrak{G}}'). \end{aligned} \tag{49}$$

In order to determine the Gaussian curvature, we have to take two arbitrary different directions $\delta_1\mathfrak{Z}$ and $\delta_2\mathfrak{Z}$ at the point $\mathfrak{Z}$ and to choose $\mathfrak{U} = \delta_1\mathfrak{Z}$, $\mathfrak{B} = \delta_2\mathfrak{Z}$. Then

$$R = -\tfrac{1}{4}\sigma(\mathfrak{Y}^{-1}\mathfrak{F}\mathfrak{Y}^{-1}\bar{\mathfrak{F}}') \leq 0,$$

where the sign of equality is true only for $\mathfrak{F} = 0$. Consequently the curvature is negative for $\delta_1\mathfrak{Z}\mathfrak{Y}^{-1}\delta_2\bar{\mathfrak{Z}} \neq \delta_2\mathfrak{Z}\mathfrak{Y}^{-1}\delta_1\bar{\mathfrak{Z}}$, and 0 otherwise.

It may easily be seen, that the contracted tensor of curvature is $\dfrac{n+1}{2}\,ds^2$; hence we have an Einstein metric with cosmological term.

18. Allendoerfer and Fenchel proved independently the following generalization of the Gauss-Bonnet formula concerning the curvatura integra of a closed two-dimensional surface. Let F be a closed manifold with the Riemann metric (43) and an even number m of dimensions. For every permutation $k_1, \cdots, k_m$ of the numbers $1, \cdots, m$ we define $\epsilon_{k_1 \ldots k_m} = 1$, if the permutation is even, and $= -1$, if the permutation is odd. Let g be the determinant $|\, g_{kl} \,|$ and

$$K = (2^{m/2} g m!)^{-1} \Sigma R_{k_1 k_2 l_1 l_2} R_{k_3 k_4 l_3 l_4} \cdots R_{k_{m-1} k_m l_{m-1} l_m} \epsilon_{k_1 k_2 \ldots k_{m-1} k_m} \epsilon_{l_1 l_2 \ldots l_{m-1} l_m}, \tag{50}$$

where the summation is extended over all permutations $k_1, \cdots, k_m$ and $l_1, \cdots, l_m$ of $1, \cdots, m$; moreover let $d\omega$ be the volume element in the given metric. Then the Euler characteristic of F has the value

$$\chi = \pi^{-(m+1)/2}\Gamma\left(\frac{m+1}{2}\right)\int_F K\,d\omega. \tag{51}$$

For practical purposes, the sum on the right-hand side of (50) may be

calculated in the following manner: Determine the single terms of the polynomial

$$R^{m/2} = (\Sigma R_{klpq} u_{ik} v_{il} u^*_{jp} v^*_{jq})^{m/2}$$

and replace every product $u_{k_1 v_{k_2}} \cdots u_{k_{m-1} v_{k_m}} u^*_{l_1} v^*_{l_2} \cdots u^*_{l_{m-1}} v^*_{l_m}$ by $\epsilon_{k_1 \ldots k_m} \epsilon_{l_1 \ldots l_m}$.

In the case of the symplectic metric, the transitive group Ω of isometric mappings exists, and consequently the invariant K is constant in the whole space H. In order to find this constant value, we may assume $\mathfrak{Z} = i\mathfrak{E}$. Writing $\mathfrak{U}^*, \mathfrak{V}^*$ instead of $\delta_1 \mathfrak{Z}, \delta_2 \mathfrak{Z}$, we obtain, by (47), (48) and (49),

$$R = \tfrac{1}{4} \sigma(\Omega \Omega^*)$$

with

$$\Omega = \mathfrak{U}\bar{\mathfrak{V}} - \mathfrak{V}\bar{\mathfrak{U}}, \qquad \Omega^* = \mathfrak{U}^* \bar{\mathfrak{V}}^* - \mathfrak{V}^* \bar{\mathfrak{U}}^*,$$

where $\mathfrak{U}, \mathfrak{V}, \mathfrak{U}^*, \mathfrak{V}^*$ are indeterminate symmetric complex matrices. Hence

$$R^\nu = 2^{-2\nu} \sum q_{k_1 l_1} \cdots q_{k_\nu l_\nu} q^*_{k_1 l_1} \cdots q^*_{k_\nu l_\nu} \qquad (\nu = 1, 2, \cdots) \tag{52}$$

with

$$q_{kl} = \sum_{r=1}^{n} (u_{kr} \bar{v}_{rl} - v_{kr} \bar{u}_{rl}), \qquad q^*_{kl} = \sum_{r=1}^{n} (u^*_{kr} \bar{v}^*_{rl} - v^*_{kr} \bar{u}^*_{rl}),$$

where $k_1, l_1, \cdots, k_\nu, l_\nu$ run independently from 1 to n. We choose

$$\nu = \frac{m}{2} = \frac{n(n+1)}{2}.$$

Let us denote the ν elements u_{kl} $(1 \leq k \leq l \leq n)$ of $\mathfrak{U} = (u_{kl})$ in lexicographic order by $u_1, \cdots u_\nu$, their conjugates by $u_{\nu+1}, \cdots, u_m$, and introduce the corresponding notation for the elements of $\mathfrak{V}, \mathfrak{U}^*, \mathfrak{V}^*$. Replacing every term $u_{g_1 v_{g_2}} \cdots u_{g_{m-1} v_{g_m}}$ in $q_{k_1 l_1} \cdots q_{k_\nu l_\nu}$ by $\epsilon_{g_1 \ldots g_m}$, we get the expression $2^\nu \eta_{k_1 l_1 \ldots k_\nu l_\nu}$ with

$$\eta_{k_1 l_1 \ldots k_\nu l_\nu} = \Sigma \epsilon_{g_1 \ldots g_m}, \tag{53}$$

where $g_1, \cdots, g_m$ run over all permutations of $1, \cdots, m$ such that the m conditions

$$u_{k_h r_h} = u_{g_{2h-1}}, \qquad \bar{v}_{l_h r_h} = v_{g_{2h}} \qquad (h = 1, \cdots, \nu)$$

have a solution $r_1, \cdots, r_\nu$. By (52), the sum of the right-hand side of (50) has, for $\mathfrak{Z} = i\mathfrak{E}$, the positive integral value

$$a_n = \Sigma \eta^2_{k_1 l_1 \ldots k_\nu l_\nu}. \tag{54}$$

Moreover g is the determinant of the quadratic form $\sigma(\mathfrak{Y}^{-1} d\mathfrak{Z} \mathfrak{Y}^{-1} d\bar{\mathfrak{Z}})$ of the variables dz_{kl} and $d\bar{z}_{kl}$ $(1 \leq k \leq l \leq n)$; hence

$$g = (-1)^\nu 2^{-2n},$$

for $\mathfrak{Z} = i\mathfrak{E}$, and consequently

$$K = (-1)^{\nu} 2^{2n-\nu} \frac{a_n}{(2\nu)!}. \tag{55}$$

On the other hand,

$$d\omega = 2^{\nu-n} \mid \mathfrak{Y} \mid^{n-1} \prod_{k \le l} (dx_{kl} dy_{kl}) = 2^{\nu-n} \prod_{k \le l} (dx_{kl} dY_{kl}) \tag{56}$$

with $(Y_{kl}) = \mathfrak{Y}^{-1}$. Since

$$\pi^{-\nu-\frac{1}{2}} \Gamma(\nu + \tfrac{1}{2}) \frac{2^{2n-\nu}}{(2\nu)!} 2^{\nu-n} = (2^{n^2} \pi^{\nu} \nu!)^{-1},$$

we obtain, by (51), (55) and (56),

$$\chi = c_n (-\pi)^{-n(n+1)/2} \int_F dv,$$

where the positive rational number

$$c_n = \frac{a_n}{2^{n^2} [(n^2 + n)/2]!}$$

is defined in (53), (54) and

$$dv = \prod_{k \le l} (dx_{kl} dY_{kl}) \tag{57}$$

denotes the euclidean volume element in the space of $\mathfrak{X}, \mathfrak{Y}^{-1}$.

We find by direct calculation $a_1 = 1$, $a_2 = 6 \cdot 3!$, $a_3 = 90 \cdot 6!$ and therefore $c_1 = \frac{1}{2}$, $c_2 = \frac{3}{8}$, $c_3 = \frac{45}{256}$, but a simple explicit formula for a_n and c_n in the case of an arbitrary n is not known.

The proof of Theorem 5 is now accomplished.

IV. DISCONTINUOUS GROUPS.

19. A group of mappings of H onto itself is called discontinuous (in H), if for every $\mathfrak{Z}$ of H the set of images of $\mathfrak{Z}$ has no limit point in H. Since H can be covered by a countable number of compact domains, the number of elements of a discontinuous group is either finite or countably infinite. We shall assume, moreover, that the mappings are analytic. Then, by Theorem 1, they form a subgroup Δ of the symplectic group.

On the other hand, let us consider the definition of a *discrete* group. A group of matrices $\mathfrak{M}$ with real (or complex) elements is called discrete, if every infinite sequence of different $\mathfrak{M}$ diverges. It is obvious that a discontinuous group of symplectic matrices is discrete. Let us now prove the con-

verse of this statement. If Δ is a non-discontinuous group of symplectic mappings

$$\mathfrak{Z}^* = (\mathfrak{A}\mathfrak{Z} + \mathfrak{B})(\mathfrak{C}\mathfrak{Z} + \mathfrak{D})^{-1}, \tag{58}$$

we can find a point $\mathfrak{Z}$ of H and an infinite sequence of different matrices

$$\mathfrak{M} = \begin{pmatrix} \mathfrak{A} & \mathfrak{B} \\ \mathfrak{C} & \mathfrak{D} \end{pmatrix}$$

in Δ, such that the corresponding sequence (58) tends to a limit $\mathfrak{Z}^*_1$ in H. Denoting by $\mathfrak{Y}, \mathfrak{Y}^*, \mathfrak{Y}^*_1$ the imaginary parts of $\mathfrak{Z}, \mathfrak{Z}^*, \mathfrak{Z}^*_1$, we have, by (25),

$$\mathfrak{Y}^* = \mathfrak{Y}\{(\mathfrak{C}\mathfrak{Z} + \mathfrak{D})^{-1}\}. \tag{59}$$

Now $\mathfrak{Y}^*$ is bounded, $\mathfrak{Y}$ is fixed and $\mathfrak{Y} > 0$. By (59), the sequence of matrices $(\mathfrak{C}\mathfrak{Z} + \mathfrak{D})^{-1}$ is bounded; moreover, the square of the absolute value of the determinant $|\mathfrak{C}\mathfrak{Z} + \mathfrak{D}|$ tends to the limit $|\mathfrak{Y}|\,|\mathfrak{Y}^*_1|^{-1}$; hence also $\mathfrak{C}\mathfrak{Z} + \mathfrak{D}$ is bounded. Since $\mathfrak{C}$ is the imaginary part of $(\mathfrak{C}\mathfrak{Z} + \mathfrak{D})\mathfrak{Y}^{-1}$ and $\mathfrak{D} = (\mathfrak{C}\mathfrak{Z} + \mathfrak{D}) - \mathfrak{C}\mathfrak{Z}$, the matrices $\mathfrak{C}, \mathfrak{D}$ are bounded. It follows from $\mathfrak{A}\mathfrak{Z} + \mathfrak{B} = \mathfrak{Z}^*(\mathfrak{C}\mathfrak{Z} + \mathfrak{D})$ that also $\mathfrak{A}$ and $\mathfrak{B}$ are bounded. Consequently there exists a converging subsequence of matrices $\mathfrak{M}$, and Δ is non-discrete.

20. For any point set P in H, we define the diameter $\delta(P)$ as the least upper bound of the distance $\rho(\mathfrak{Z}, \mathfrak{Z}^*)$, where $\mathfrak{Z}$ and $\mathfrak{Z}^*$ run independently over all points of P. The diameter is finite, if P is compact. The distance $\rho(P, P^*)$ of two point sets P, P^* in H is defined as the greatest lower bound of the distance $\rho(\mathfrak{Z}, \mathfrak{Z}^*)$, $\mathfrak{Z}$ running over P and $\mathfrak{Z}^*$ over P^*. By the triangle inequality,

$$\delta(P + P^*) \leq \rho(P, P^*) + \delta(P) + \delta(P^*). \tag{60}$$

Let $D_1, D_2, \cdots$ be the elements of the discontinuous group Δ and let D_1 be the identity. We denote by $D_k(P) = P_k$ the image of the set P under D_k $(k = 1, 2, \cdots)$. We assume now that P and another point set Q in H are compact. We shall prove that the distance $\rho(Q, P_k)$ tends to infinity with k. Let us first consider the case of two points $Q = \mathfrak{Z}$, $P = \mathfrak{Z}_1$. By Section **16**, the condition $\rho(\mathfrak{Z}, \mathfrak{Z}^*) \leq c$ defines, for arbitrarily given $c > 0$ and variable $\mathfrak{Z}^*$, a compact point set in H; hence the inequality $\rho(\mathfrak{Z}, \mathfrak{Z}_k) \leq c$ holds only for a finite number of indices k so that $\rho(\mathfrak{Z}, \mathfrak{Z}_k) \to \infty$ for $k \to \infty$. Consider now the general case for P and Q. By (60),

$$\rho(Q, P_k) \geq \delta(Q + P_k) - \delta(Q) - \delta(P_k).$$

Choosing a point $\mathfrak{Z}$ in Q and a point $\mathfrak{Z}_1$ in P, we have $\delta(Q + P_k) \geq \rho(\mathfrak{Z}, \mathfrak{Z}_k) \to \infty$ and $\delta(P_k) = \delta(P)$, whence $\rho(Q, P_k) \to \infty$.

A point $\mathfrak{Z}$ of H is called a fixed point of Δ, if $D_k(\mathfrak{Z}) = \mathfrak{Z}$ for at least one index $k > 1$. Let P be an arbitrary compact domain in H, e. g., the domain $\rho(\mathfrak{Z}, i\mathfrak{E}) \leq 1$. Since $\rho(P, P_k) \to \infty$ for $k \to \infty$, there exists only a finite number of values k, such that the equation $D_k(\mathfrak{Z}) = \mathfrak{Z}$ has a solution $\mathfrak{Z}$ in P. But this equation, for any $k > 1$, defines an algebraic manifold in H, and consequently we may construct a point of P which is not contained in any of those manifolds; in other words, we may certainly construct a point $\mathfrak{Z}_1$ of H which is not a fixed point. The images $\mathfrak{Z}_k = D_k(\mathfrak{Z}_1)$ of $\mathfrak{Z}_1$ are then all different one from another.

21. We denote by F the set of all points $\mathfrak{Z}$ satisfying all the inequalities

$$\rho(\mathfrak{Z}, \mathfrak{Z}_1) \leq \rho(\mathfrak{Z}, \mathfrak{Z}_k) \qquad (k = 2, 3, \cdots). \tag{61}$$

It follows from this definition that F is closed, with respect to H; but F is not necessarily compact. Let $G = H - F$ be the complement of G in H and let B be the frontier of F and $F_0 = F - B$ the set of inner points of F.

Obviously, the set G consists of all points $\mathfrak{Z}$ satisfying the inequality $\rho(\mathfrak{Z}, \mathfrak{Z}_1) > \rho(\mathfrak{Z}, \mathfrak{Z}_k)$ for at least one value of k; hence all the points $\mathfrak{Z}_2, \mathfrak{Z}_3, \cdots$ belong to G.

Let us now consider a point $\mathfrak{Z}$ which fulfills all conditions

$$\rho(\mathfrak{Z}, \mathfrak{Z}_1) < \rho(\mathfrak{Z}, \mathfrak{Z}_k) \qquad (k = 2, 3, \cdots); \tag{62}$$

the point $\mathfrak{Z} = \mathfrak{Z}_1$ is an example. The differences $\rho(\mathfrak{Z}, \mathfrak{Z}_k) - \rho(\mathfrak{Z}, \mathfrak{Z}_1)$ $(k = 2, 3, \cdots)$ are all positive and tend to infinity with k and, consequently, they have a positive minimum μ. The points $\mathfrak{Z}^*$ of the geodesic sphere $\rho(\mathfrak{Z}, \mathfrak{Z}^*) < \frac{1}{2}\mu$ form a neighborhood of $\mathfrak{Z}$. It follows from

$$\rho(\mathfrak{Z}^*, \mathfrak{Z}_k) - \rho(\mathfrak{Z}^*, \mathfrak{Z}_1) > \rho(\mathfrak{Z}, \mathfrak{Z}_k) - \tfrac{1}{2}\mu - \rho(\mathfrak{Z}, \mathfrak{Z}_1) - \tfrac{1}{2}\mu \geq 0 \qquad (k = 2, 3, \cdots),$$

that all these points belong to F. Consequently $\mathfrak{Z}$ is a point of F_0.

Consider next the case where all conditions (61) are fulfilled, with the sign of equality for at least one index $k = l > 1$. Then $\mathfrak{Z} \neq \mathfrak{Z}_l$. Let $\mathfrak{Z}^*$ be an arbitrary point on the geodesic arc joining $\mathfrak{Z}$ and $\mathfrak{Z}_l$, different from $\mathfrak{Z}$. Since $\rho(\mathfrak{Z}, \mathfrak{Z}_1) = \rho(\mathfrak{Z}, \mathfrak{Z}_l)$ and $\mathfrak{Z}_1 \neq \mathfrak{Z}_l$, the point $\mathfrak{Z}^*$ does not lie at the same time on the geodesic arc between $\mathfrak{Z}$ and $\mathfrak{Z}_1$ so that

$$\rho(\mathfrak{Z}, \mathfrak{Z}_1) < \rho(\mathfrak{Z}, \mathfrak{Z}^*) + \rho(\mathfrak{Z}^*, \mathfrak{Z}_1).$$

On the other hand,

$$\rho(\mathfrak{Z}, \mathfrak{Z}_l) = \rho(\mathfrak{Z}, \mathfrak{Z}^*) + \rho(\mathfrak{Z}^*, \mathfrak{Z}_l),$$

and the inequality $\rho(\mathfrak{Z}^*, \mathfrak{Z}_1) > \rho(\mathfrak{Z}^*, \mathfrak{Z}_l)$ is proved. Consequently, the whole

geodesic arc between $\mathfrak{Z}$ and $\mathfrak{Z}_l$ belongs to G, except the point $\mathfrak{Z}$ itself, which is a point of F. This proves that $\mathfrak{Z}$ is a point of B. Therefore F_0 consists of all $\mathfrak{Z}$ satisfying (62) and B consists of all $\mathfrak{Z}$ satisfying (61) with at least one sign of equality.

Let again $\mathfrak{Z}$ be a point of B, and choose a point $\mathfrak{Z}^* \neq \mathfrak{Z}$ on the geodesic arc joining $\mathfrak{Z}$ and $\mathfrak{Z}_1$; then

$$\rho(\mathfrak{Z}, \mathfrak{Z}_1) = \rho(\mathfrak{Z}, \mathfrak{Z}^*) + \rho(\mathfrak{Z}^*, \mathfrak{Z}_1).$$

Moreover

$$\rho(\mathfrak{Z}, \mathfrak{Z}_1) \leq \rho(\mathfrak{Z}, \mathfrak{Z}_k) \leq \rho(\mathfrak{Z}, \mathfrak{Z}^*) + \rho(\mathfrak{Z}^*, \mathfrak{Z}_k) \qquad (k = 2, 3, \cdots),$$

where the sign of equality cannot be true in both places. Hence $\rho(\mathfrak{Z}^*, \mathfrak{Z}_1) < \rho(\mathfrak{Z}^*, \mathfrak{Z}_k)$, i. e., $\mathfrak{Z}^*$ is a point of F_0. This proves that any geodesic ray through $\mathfrak{Z}_1$ either lies completely in F or intersects B in exactly one point. The domain F is a star formed by geodesic arcs through $\mathfrak{Z}_1$.

The boundary B of F consists of parts of the analytic surfaces $\rho(\mathfrak{Z}, \mathfrak{Z}_1) = \rho(\mathfrak{Z}, \mathfrak{Z}_k)$ for certain values of $k > 1$. It is not generally true that the number of these values is finite. However, if we consider a compact domain P in H, the distance $\rho(P, \mathfrak{Z}_k)$ tends to infinity with k; consequently only a finite number of these bounding surfaces enter into P.

Let $\mathfrak{Z}$ be an arbitrary point of H. Since $\rho(\mathfrak{Z}, \mathfrak{Z}_k) \to \infty$, there exists a positive integer r, such that

$$\rho(\mathfrak{Z}, \mathfrak{Z}_r) \leq \rho(\mathfrak{Z}, \mathfrak{Z}_k) \qquad (k = 1, 2, \cdots). \tag{63}$$

Then the point $\mathfrak{Z}^* = D_r^{-1}(\mathfrak{Z})$ satisfies the conditions (61). Consequently $\mathfrak{Z}$ is equivalent to a point of F. On the other hand, a point $\mathfrak{Z}$ cannot satisfy both conditions (62) and (63), for any $r > 1$ and all k. It follows that no point of F_0 is equivalent to any point of F, except to itself under the identical mapping D_1.

Our results contain the proof of Theorem 6.

22. If the fundamental domain F is compact, then the boundary B consists of a finite number of surfaces $\rho(\mathfrak{Z}, \mathfrak{Z}_1) = \rho(\mathfrak{Z}, \mathfrak{Z}_k)$. Now F depends upon the initial point $\mathfrak{Z}_1$; we write more explicitly $F = F(\mathfrak{Z}_1)$. We shall prove that $F(\mathfrak{Z}^*_1)$ is compact, for an arbitrary initial point $\mathfrak{Z}^*_1$, if $F(\mathfrak{Z}_1)$ is compact.

The space H is called compact relative to Δ, if there exists for every infinite sequence of points $\mathfrak{Z}^{(k)}$, $(k = 1, 2, \cdots)$, in H at least one compact sequence of images $\mathfrak{Z}_{l_k}^{(k)}$ under Δ. Obviously this condition is satisfied, if and only if there exists a compact domain G in H, such that every point $\mathfrak{Z}$

of H has at least one image $\mathfrak{Z}_k$ in G. By the minimum property of F, this domain is then also compact, and vice versa. Hence our assertion is proved.

In the classical case $n = 1$, the fundamental polygon F is compact, if and only if all vertices are elliptic; it is well-known that the uniformization of any field of algebraic functions of genus $p > 1$ leads to a discontinuous group with the required property. From the algebraic point of view, the most important discontinuous groups Δ, in the case $n = 1$, are, more generally, those having a fundamental polygon with a finite number of elliptic or parabolic vertices. They constitute the Fuchsian groups of the first kind, and the corresponding automorphic functions form algebraic function fields of a single variable.

For arbitrary n, we say that a discontinuous group Δ is of the first kind, if there exists a fundamental domain F with the following three properties: 1) Every compact domain in H is covered by a finite number of images of F; 2) only a finite number of images of F are neighbors of F; 3) the integral

$$V(\Delta) = \int_F dv \tag{64}$$

converges. In the special case $n = 1$, it is easily seen that this definition is tantamount to the ordinary definition of Fuchsian groups of the first kind.

It is now clear that Δ is certainly of the first kind, if H is compact relative to Δ; hence Theorem 7 is proved.

Let

$$\mathfrak{M} = \begin{pmatrix} \mathfrak{A} & \mathfrak{B} \\ \mathfrak{C} & \mathfrak{D} \end{pmatrix}$$

be any symplectic matrix. Under the automorphism $\mathfrak{Z} \to (\mathfrak{A}\mathfrak{Z} + \mathfrak{B})(\mathfrak{C}\mathfrak{Z} + \mathfrak{D})^{-1}$ of H, a subgroup Δ of Ω is replaced by $\mathfrak{M}\Delta\mathfrak{M}^{-1}$, i. e., by a conjugate subgroup, and a fundamental domain of Δ is mapped onto a fundamental domain of $\mathfrak{M}\Delta\mathfrak{M}^{-1}$. Obviously all conjugate subgroups $\mathfrak{M}\Delta\mathfrak{M}^{-1}$ will be of the first kind, if Δ itself is of the first kind.

23. Let us now assume that Δ has no fixed point in H. Then the images $\mathfrak{Z}_k (k = 1, 2, \cdots)$ of an arbitrary point $\mathfrak{Z}_1$ of H are all different from each other, and the minimum of the distances $\rho(\mathfrak{Z}_k, \mathfrak{Z}_1)$ $(k = 2, 3, \cdots)$ is a positive number $\delta = \delta(\mathfrak{Z}_1)$. The images of the geodesic sphere $\rho(\mathfrak{Z}, \mathfrak{Z}_1) < \frac{1}{2}\delta$ with the center $\mathfrak{Z}_1$ do not overlap.

Identifying equivalent points of H, we obtain a set H_Δ. We may obviously introduce the symplectic metric into H_Δ, defining a neighborhood of $\mathfrak{Z}_1$ by $\rho(\mathfrak{Z}, \mathfrak{Z}_1) < \frac{1}{2}\delta$. Starting from a fundamental domain F of Δ, we obtain a

model of H_Δ, if we identify equivalent points on the boundary of F. Assume now that H is compact relative to Δ. Then F may be considered as a closed manifold with the symplectic metric. By Theorem 5, this manifold has the Euler number

$$\chi = c_n(-\pi)^{-n(n+1)/2}V(\Delta), \tag{65}$$

where $V(\Delta)$ denotes the symplectic volume (6) of F.

Probably the formula (65) is true for all groups Δ of the first kind, provided Δ has no fixed point in H. This is easily proved in the case $n = 1$. The general proof of our suggestion would require a careful investigation of the geodesics of infinite length in the fundamental domain.

V. HERMITIAN FORMS.

24. In the case $n = 1$, we know three different methods of constructing discontinuous groups of the first kind, namely an analytic, a geometric and an arithmetic method. The analytic method starts with a Riemann surface of finite genus and applies the theory of uniformization. The geometric method uses the principle of reflection for a circular polygon with a finite number of elliptic and parabolic vertices, the angles at the elliptic vertices being aliquot parts of π. The arithmetic method depends upon the theory of units of indefinite binary hermitian forms, in an imaginary quadratic ring over a totally real algebraic number field of finite degree. It is not known to what extent the analytic and the geometric method may be generalized; however, we shall show in the following sections, that there is a generalization of the arithmetic method to the case of an arbitrary n.

LEMMA 3. *Let $\mathfrak{H}^{(2n)}$ be a hermitian and $\mathfrak{G}^{(2n)}$ a non-singular skew-symmetric matrix with complex elements. If*

$$\mathfrak{H}\bar{\mathfrak{G}}^{-1}\bar{\mathfrak{H}} = \mathfrak{G}, \tag{66}$$

then there exists a matrix $\mathfrak{C}$ such that $\mathfrak{H}\{\mathfrak{C}\} = i\mathfrak{J}$ and $\mathfrak{G}[\mathfrak{C}] = \mathfrak{J}$.

Putting $\mathfrak{G}^{-1}\mathfrak{H} = \mathfrak{F}$, we have $\mathfrak{F}\bar{\mathfrak{F}} = \mathfrak{E}$, by (66), whence

$$(\mathfrak{F} + i\lambda\mathfrak{E})(\bar{\mathfrak{F}} + i\bar{\lambda}\mathfrak{E}) + (\mathfrak{F} - i\lambda\mathfrak{E})(\bar{\mathfrak{F}} - i\bar{\lambda}\mathfrak{E}) = 0$$

for any scalar λ of absolute value 1. Choosing this λ such that $|\mathfrak{F} + i\lambda\mathfrak{E}| \neq 0$, we obtain

$$\frac{\mathfrak{F} - i\lambda\mathfrak{E}}{\mathfrak{F} + i\lambda\mathfrak{E}} = i\mathfrak{T}$$

with real $\mathfrak{T}$; consequently $\mathfrak{E} - i\mathfrak{T} = 2i\lambda(\mathfrak{F} + i\lambda\mathfrak{E})^{-1}$, $|\mathfrak{E} - i\mathfrak{T}| \neq 0$,

$$\mathfrak{F} = i\lambda \frac{\mathfrak{E} + i\mathfrak{T}}{\mathfrak{E} - i\mathfrak{T}}.$$

Let $\mathfrak{C}_1 = \lambda^{\frac{1}{2}}(\mathfrak{E} + i\mathfrak{T})$, $\mathfrak{H}_1 = \mathfrak{H}\{\mathfrak{C}_1\}$, $\mathfrak{G}_1 = \mathfrak{G}[\mathfrak{C}_1]$; then

$$\mathfrak{G}_1^{-1}\mathfrak{H}_1 = \mathfrak{C}_1^{-1}\mathfrak{F}\bar{\mathfrak{C}}_1 = i\mathfrak{E}; \quad \bar{\mathfrak{H}}_1 = \mathfrak{H}'_1 = i\mathfrak{G}'_1 = -i\mathfrak{G}_1 = -\mathfrak{H}_1;$$

hence $\mathfrak{H}_1$ is pure imaginary and $\mathfrak{G}_1 = -i\mathfrak{H}$ real. There exists a real matrix $\mathfrak{C}_2$ satisfying $\mathfrak{G}_1[\mathfrak{C}_2] = \mathfrak{J}$, and $\mathfrak{C} = \mathfrak{C}_1\mathfrak{C}_2$ has the required property.

25. Let K be a totally real algebraic number field of finite degree h and let $K^{(1)}, \cdots, K^{(h)}$ be its conjugate fields, $K = K^{(1)}$. If r is any positive number of K, then the field $K_0 = K(\sqrt{-r})$, of degree $2h$ is imaginary. We consider a hermitian matrix $\mathfrak{H} = \mathfrak{H}^{(2n)}$ and a non-singular skew-symmetric matrix $\mathfrak{G}$, both with elements in K_0, and we assume that the relationship

$$\mathfrak{H}\bar{\mathfrak{G}}^{-1}\bar{\mathfrak{H}} = s\mathfrak{G} \tag{67}$$

holds with a positive scalar factor s. Obviously s is then a number of K.

The matrices $\mathfrak{U}$ with integral elements in K_0, satisfying the two conditions

$$\mathfrak{G}[\mathfrak{U}] = \mathfrak{G}, \qquad \mathfrak{H}\{\mathfrak{U}\} = \mathfrak{H},$$

constitute a multiplicative group $\Lambda = \Lambda(\mathfrak{G}, \mathfrak{H})$. Applying Lemma 3 with $s^{-\frac{1}{2}}\mathfrak{H}$ instead of $\mathfrak{H}$, we obtain a complex matrix $\mathfrak{C}$, such that $\mathfrak{G}[\mathfrak{C}] = \mathfrak{J}$ and $\mathfrak{H}\{\mathfrak{C}\} = is^{\frac{1}{2}}\mathfrak{J}$. Consequently the elements $\mathfrak{C}^{-1}\mathfrak{U}\mathfrak{C} = \mathfrak{M}$ of the group $\mathfrak{C}^{-1}\Lambda\mathfrak{C} = \Delta_0 = \Delta_0(\mathfrak{G}, \mathfrak{H})$ satisfy $\mathfrak{J}[\mathfrak{M}] = \mathfrak{J}$ and $\mathfrak{J}\{\mathfrak{M}\} = \mathfrak{J}$, whence $\mathfrak{J}[\mathfrak{M}] = \mathfrak{J}$ and $\bar{\mathfrak{M}} = \mathfrak{M}$. This proves that Δ_0 is a subgroup of the homogeneous symplectic group Ω_0. Identifying $\mathfrak{U}$ and $-\mathfrak{U}$, i. e., $\mathfrak{M}$ and $-\mathfrak{M}$, we obtain a subgroup $\Delta = \Delta(\mathfrak{G}, \mathfrak{H})$ of Ω.

The matrix $\mathfrak{C}$ is not uniquely determined. If also $\mathfrak{H}\{\mathfrak{C}^*\} = is^{\frac{1}{2}}\mathfrak{J}$ and $\mathfrak{G}[\mathfrak{C}^*] = \mathfrak{J}$, then $\mathfrak{C}^{-1}\mathfrak{C}^* = \mathfrak{B}$ is symplectic, and vice versa. Using $\mathfrak{C}^*$ instead of $\mathfrak{C}$, we have to replace Δ by $\mathfrak{B}^{-1}\Delta\mathfrak{B}$; hence the class of conjugate subgroups $\mathfrak{B}^{-1}\Delta\mathfrak{B}$ in Ω is uniquely determined by $\mathfrak{G}$ and $\mathfrak{H}$.

Obviously $\Lambda(\mathfrak{G}, a\mathfrak{H}) = \Lambda(\mathfrak{G}, \mathfrak{H})$ for any number $a \neq 0$ of K. Therefore we may assume r and s to be integers. Henceforth we shall, moreover, assume that r is totally positive and that all conjugates of $\mathfrak{H}$ except $\mathfrak{H}$ and $\bar{\mathfrak{H}}$ are positive. Let r_l and the pair $\mathfrak{H}_l, \bar{\mathfrak{H}}_l$ be the image of r and the pair $\mathfrak{H}, \bar{\mathfrak{H}}$ under the isomorphism $K \to K^{(l)}$ $(l = 1, \cdots, h)$. For any element $\mathfrak{U}$ of Λ, we have $\mathfrak{H}_l\{\mathfrak{U}_l\} = \mathfrak{H}_l$, where $\mathfrak{U}_l, \bar{\mathfrak{U}}_l$ $(l = 1, \cdots, h)$ denote the $2h$ conjugates of $\mathfrak{U} = \mathfrak{U}_1$. Since $\mathfrak{H}_l$ is positive for $l > 1$, the matrices $\mathfrak{U}_2, \cdots, \mathfrak{U}_h$ are bounded. If also $\mathfrak{U}$ itself is bounded, then all conjugates of $\mathfrak{U}$ are bounded.

On the other hand, there exists only a finite number of integers in K_0 with bounded conjugates. Consequently Λ and Δ_0 are discrete groups and, by Section **19**, the group $\Delta(\mathfrak{G}, \mathfrak{H})$ is discontinuous.

It follows from (67) that $\mathfrak{H} = -s\bar{\mathfrak{H}}^{-1}\{\mathfrak{G}\}$. Since s is positive, $\mathfrak{H} = \mathfrak{H}_1$ is necessarily the matrix of an indefinite hermitian form. Since $\mathfrak{H}_l$ is positive for $l > 1$, the conjugates $s_2, \cdots, s_h$ of the positive number $s = s_1$ are all negative.

The most important example of a group $\Delta(\mathfrak{G}, \mathfrak{H})$ is provided by $\mathfrak{G} = \mathfrak{J}$, $\mathfrak{H} = i\mathfrak{J}$, $r = 1$, $h = 1$. Then we may choose $\mathfrak{C} = \mathfrak{E}$, and $\Delta_0(\mathfrak{G}, \mathfrak{H})$ consists of all symplectic matrices $\mathfrak{M}$ with rational integral elements. We call this group the homogeneous modular group of degree n and denote it by Γ_0. Identifying the elements $\mathfrak{M}$ and $-\mathfrak{M}$ of Γ_0, we obtain the (inhomogeneous) modular group Γ.

26. Two subgroups Δ and Δ^* of Ω are called commensurable, if there exist a subgroup Δ_1 of finite index in Δ and a subgroup Δ^*_1 of finite index in Δ^*, such that Δ_1 and Δ^*_1 are conjugate subgroups of Ω. If Δ is a discontinuous group of the first kind, then the same holds for Δ^*, and we obtain $jV(\Delta) = j^*V(\Delta^*)$, where j and j^* denote the indices of the subgroups Δ_1 and Δ^*_1; consequently the quotient $V(\Delta)/V(\Delta^*)$ is a rational number.

It is easily seen that the property of commensurability is symmetric and transitive; therefore we may speak of a class of commensurable groups. We have now the problem of deciding whether two groups $\Delta(\mathfrak{G}, \mathfrak{H})$ and $\Delta(\mathfrak{G}^*, \mathfrak{H}^*)$ are commensurable or not. The complete answer is given by Theorem 13. In this section we solve only a particular case of the problem: We assume that $\mathfrak{G}^*$ and $\mathfrak{H}^*$ are also matrices of the field K_0 and that they fulfill the condition

$$\mathfrak{H}^*\bar{\mathfrak{G}}^{*-1}\bar{\mathfrak{H}}^* = s\mathfrak{G}^*$$

with the same factor s as in (67).

LEMMA 4. *Let $c_1, \cdots, c_{2n}$ be $2n$ numbers of K_0, not all 0. There exists a matrix $\mathfrak{C}_1^{(2n)} = (c_{kl})$ in K_0, such that $c_{k1} = c_k$ $(k = 1, \cdots, 2n)$ and $\mathfrak{J}[\mathfrak{C}_1] = \mathfrak{J}$. If, moreover, $c_1 \neq 0$, we may choose $c_{1l} = 0$ $(l = 2, \cdots, 2n)$.*

Put $\mathfrak{c}'_1 = (c_1 \cdots c_n)$ and $\mathfrak{c}'_2 = (c_{n+1} \cdots c_{2n})$. If $\mathfrak{c}_1 = 0$, we choose in K_0 a non-singular matrix $\mathfrak{Q}^{(n)}$ with the first row $\mathfrak{c}'_2$ and define

$$\mathfrak{C}_1 = \begin{pmatrix} 0 & -\mathfrak{Q}^{-1} \\ \mathfrak{Q}' & 0 \end{pmatrix}.$$

If $\mathfrak{c}_1 \neq 0$, we choose in K_0 a non-singular matrix $\mathfrak{P}^{(n)}$ with the first row $\mathfrak{c}'_1$ and a symmetric matrix $\mathfrak{S}$ with the first column $\mathfrak{P}\mathfrak{c}_2$; then

$$\mathfrak{C}_1 = \begin{pmatrix} \mathfrak{P}' & 0 \\ \mathfrak{P}^{-1}\mathfrak{S} & \mathfrak{P}^{-1} \end{pmatrix}$$

has the required property. In the case $c_1 \neq 0$, we may obviously choose $(c_1 0 \cdots 0)$ as the first row of $\mathfrak{P}'$.

LEMMA 5. *There exists a matrix* $\mathfrak{C}_0$ *in* K_0 *such that* $\mathfrak{G}[\mathfrak{C}_0] = \mathfrak{G}^*$ *and* $\mathfrak{H}\{\mathfrak{C}_0\} = \mathfrak{H}^*$.

The equation $\mathfrak{G}[\mathfrak{X}] = \mathfrak{J}$ has a solution $\mathfrak{X}$ in K_0; hence we may assume, without loss of generality, that $\mathfrak{G} = \mathfrak{G}^* = \mathfrak{J}$. Moreover it is sufficient to prove the lemma for the special case

$$\mathfrak{H}^* = \begin{pmatrix} \mathfrak{E} & 0 \\ 0 & -s\mathfrak{E} \end{pmatrix}.$$

Since none of the conjugates of $-\mathfrak{H}$ is positive, it follows from Hasse's generalization of A. Meyer's theorem, that the diophantine equation $\mathfrak{H}\{\mathfrak{x}\} = 1$ has a solution $\mathfrak{x}$ in K_0. Applying Lemma 4 with $(c_1 \cdots c_{2n}) = \mathfrak{x}'$, we construct a matrix $\mathfrak{C}_1$ in K_0 which has the first column $\mathfrak{x}$ and satisfies $\mathfrak{J}[\mathfrak{C}_1] = \mathfrak{J}$. Then $\mathfrak{H}\{\mathfrak{C}_1\}$ has the first diagonal element 1. On the other hand, the condition (67) means now

$$\mathfrak{H}\mathfrak{J}\bar{\mathfrak{H}} = -s\mathfrak{J}. \tag{68}$$

For the proof of Lemma 5, we may therefore replace $\mathfrak{H}\{\mathfrak{C}_1\}$ again by $\mathfrak{H}$.

By Lemma 4, there exists a matrix $\mathfrak{C}_2$ in K_0 with the following three properties: $\mathfrak{C}_2$ and $\mathfrak{H}$ have the same first column; $\mathfrak{C}_2$ has the first row $(1\,0 \cdots 0)$; $\mathfrak{J}[\mathfrak{C}_2] = \mathfrak{J}$. Put $\mathfrak{C}_2^{-1} = \mathfrak{C}'_3$; then also $\mathfrak{J}[\mathfrak{C}_3] = \mathfrak{J}$, and the two matrices $\mathfrak{C}'_3\mathfrak{H}$ and $\bar{\mathfrak{C}}_3$ have both the first column $(1\,0 \cdots 0)'$. Writing again $\mathfrak{H}$ instead of $\mathfrak{H}\{\mathfrak{C}_3\}$, we obtain the decomposition

$$\mathfrak{H} = \begin{pmatrix} \mathfrak{H}_1 & \mathfrak{H}_{12} \\ \bar{\mathfrak{H}}'_{12} & \mathfrak{H}_2 \end{pmatrix},$$

$$\mathfrak{H}_1{}^{(n)} = \begin{pmatrix} 1 & 0 \\ 0 & \mathfrak{F}_1 \end{pmatrix}, \qquad \mathfrak{H}_{12} = \begin{pmatrix} 0 & 0 \\ \mathfrak{a} & \mathfrak{F}_{12} \end{pmatrix}, \qquad \mathfrak{H}_2 = \begin{pmatrix} t & \mathfrak{c}' \\ \mathfrak{b} & \mathfrak{F}_2 \end{pmatrix}.$$

By (68),

$$\mathfrak{H}_1\mathfrak{H}'_{12} = \mathfrak{H}_{12}\mathfrak{H}'_1, \qquad \bar{\mathfrak{H}}'_{12}\mathfrak{H}'_2 = \mathfrak{H}_2\bar{\mathfrak{H}}_{12},$$

$$\mathfrak{H}_1\mathfrak{H}'_2 - \mathfrak{H}_{12}\bar{\mathfrak{H}}_{12} = -s\mathfrak{E}, \qquad \mathfrak{H}'_1\mathfrak{H}_2 - \bar{\mathfrak{H}}_{12}\mathfrak{H}_{12} = -s\mathfrak{E},$$

whence $\mathfrak{a} = 0$, $t = -s$, $\mathfrak{b} = 0$, $\mathfrak{c} = 0$. This contains, in particular, the proof of the lemma for $n = 1$. If $n > 1$, the hermitian matrix

$$\mathfrak{F} = \begin{pmatrix} \mathfrak{F}_1 & \mathfrak{F}_{12} \\ \bar{\mathfrak{F}}'_{12} & \mathfrak{F}_2 \end{pmatrix}$$

fulfills the same conditions as $\mathfrak{H}$, with $n-1$ instead of n; and we may apply induction with respect to n. Hence the lemma is proved.

Let $\mathfrak{C}_0$ be the matrix of Lemma 5 and choose a positive rational integer q, such that the two matrices $q\mathfrak{C}_0$ and $q\mathfrak{C}_0^{-1}$ have integral elements. Since the integers of K_0 belong to a finite number of classes of residues modulo q^2, the elements $\mathfrak{U}$ of $\Lambda(\mathfrak{G},\mathfrak{H})$ satisfying $\mathfrak{U}\equiv\mathfrak{E} \pmod{q^2}$ form a subgroup Λ_1 of finite index. Consider now the subgroup Λ_2 consisting of all $\mathfrak{U}$ with integral $\mathfrak{C}_0^{-1}\mathfrak{U}\mathfrak{C}_0=\mathfrak{U}^*$. Obviously Λ_1 is contained in Λ_2; consequently Λ_2 is a fortiori of finite index in $\Lambda(\mathfrak{G},\mathfrak{H})$. On the other hand, $\mathfrak{U}^*$ is an element of $\Lambda(\mathfrak{G}^*,\mathfrak{H}^*)$ with the characteristic property that $\mathfrak{C}_0\mathfrak{U}^*\mathfrak{C}_0^{-1}=\mathfrak{U}$ is integral; hence the group $\Lambda^*{}_2=\mathfrak{C}_0^{-1}\Lambda_2\mathfrak{C}_0$ is of finite index in $\Lambda(\mathfrak{G}^*,\mathfrak{H}^*)$. If $\mathfrak{C}$ is the matrix of Section **25**, we have $\Delta_0(\mathfrak{G},\mathfrak{H})=\mathfrak{C}^{-1}\Lambda(\mathfrak{G},\mathfrak{H})\mathfrak{C}$, and we may define $\Delta_0(\mathfrak{G}^*,\mathfrak{H}^*)=\mathfrak{C}_4^{-1}\Lambda(\mathfrak{G}^*,\mathfrak{H}^*)\mathfrak{C}_4$ with $\mathfrak{C}_4=\mathfrak{C}_0^{-1}\mathfrak{C}$. Then $\mathfrak{C}^{-1}\Lambda_2\mathfrak{C}=\mathfrak{C}_4^{-1}\Lambda^*{}_2\mathfrak{C}_4$ is a common subgroup of $\Delta_0(\mathfrak{G},\mathfrak{H})$ and $\Delta_0(\mathfrak{G}^*,\mathfrak{H}^*)$, of finite indices. This proves that $\Delta(\mathfrak{G},\mathfrak{H})$ and $\Delta(\mathfrak{G}^*,\mathfrak{H}^*)$ are commensurable.

27. The two conditions $\mathfrak{G}[\mathfrak{U}]=\mathfrak{G}$ and $\mathfrak{H}\{\mathfrak{U}\}=\mathfrak{H}$ for the elements $\mathfrak{U}$ of $\Lambda(\mathfrak{G},\mathfrak{H})$ may be written $\mathfrak{F}\bar{\mathfrak{U}}=\mathfrak{U}\mathfrak{F}$ and $\mathfrak{U}\mathfrak{G}^{-1}\mathfrak{U}'\mathfrak{G}=\mathfrak{E}$, with $\mathfrak{F}=\mathfrak{G}^{-1}\mathfrak{H}$. Let us consider the set R of all matrices $\mathfrak{V}$ in K_0 which satisfy the condition

$$\mathfrak{F}\bar{\mathfrak{V}}=\mathfrak{V}\mathfrak{F}. \tag{69}$$

Obviously R is a ring. By (67), the matrix

$$\bar{\mathfrak{V}}=\mathfrak{G}^{-1}\mathfrak{V}'\mathfrak{G} \tag{70}$$

is again a solution of (69), and consequently (70) defines an anti-automorphism of R. The elements $\mathfrak{U}$ of $\Lambda(\mathfrak{G},\mathfrak{H})$ are the integral elements $\mathfrak{V}=\mathfrak{U}$ of R with the property

$$\mathfrak{U}\bar{\mathfrak{U}}=\mathfrak{E}. \tag{71}$$

By Lemma 5,

$$\mathfrak{G}[\mathfrak{C}_0]=\mathfrak{J},\quad \mathfrak{H}\{\mathfrak{C}_0\}=\begin{pmatrix}\mathfrak{E} & 0\\ 0 & -s\mathfrak{E}\end{pmatrix}=\mathfrak{H}^*,\quad \mathfrak{C}_0^{-1}\mathfrak{F}\mathfrak{C}_0=\begin{pmatrix}0 & s\mathfrak{E}\\ \mathfrak{E} & 0\end{pmatrix}=\mathfrak{F}^*,$$

where $\mathfrak{C}_0$ is a matrix in K_0. The elements $\mathfrak{V}^*=\mathfrak{C}_0^{-1}\mathfrak{V}\mathfrak{C}_0$ of the ring $R^*=\mathfrak{C}_0^{-1}R\mathfrak{C}_0$ satisfy $\mathfrak{F}^*\bar{\mathfrak{V}}^*=\mathfrak{V}^*\mathfrak{F}^*$, whence

$$\mathfrak{V}^*=\begin{pmatrix}\mathfrak{P} & s\bar{\mathfrak{Q}}\\ \mathfrak{Q} & \bar{\mathfrak{P}}\end{pmatrix}$$

with arbitrary matrices $\mathfrak{P}^{(n)}$, $\mathfrak{Q}^{(n)}$ in K_0. Defining

$$2^{1/2}s^{1/4}\mathfrak{Q}=\begin{pmatrix}is^{1/2}\mathfrak{E} & s^{1/2}\mathfrak{E}\\ -\mathfrak{E} & -i\mathfrak{E}\end{pmatrix},\qquad \mathfrak{C}_0\mathfrak{Q}=\mathfrak{C},$$

we obtain

$$\mathfrak{G}[\mathfrak{C}] = \mathfrak{J}, \qquad \mathfrak{H}\{\mathfrak{C}\} = is^{\frac{1}{2}}\mathfrak{J}, \tag{72}$$

$$\mathfrak{C}^{-1}\mathfrak{B}\mathfrak{C} = \mathfrak{L}^{-1}\mathfrak{B}^{*}\mathfrak{L} = \begin{pmatrix} \mathfrak{A}_0 + \sqrt{r}s\mathfrak{A}_1 & \sqrt{r}\mathfrak{A}_2 + \sqrt{s}\mathfrak{A}_3 \\ -\sqrt{r}\mathfrak{A}_2 + \sqrt{s}\mathfrak{A}_3 & \mathfrak{A}_0 - \sqrt{r}s\mathfrak{A}_1 \end{pmatrix} = \mathfrak{M} \tag{73}$$

with

$$\mathfrak{A}_0 = \tfrac{1}{2}(\mathfrak{P} + \bar{\mathfrak{P}}), \quad \mathfrak{A}_1 = \frac{1}{2\sqrt{-r}}(\mathfrak{Q} - \bar{\mathfrak{Q}}),$$

$$\mathfrak{A}_2 = \frac{1}{2\sqrt{-r}}(\mathfrak{P} - \bar{\mathfrak{P}}), \quad \mathfrak{A}_3 = -\tfrac{1}{2}(\mathfrak{Q} + \bar{\mathfrak{Q}});$$

consequently $\mathfrak{A}_0, \mathfrak{A}_1, \mathfrak{A}_2, \mathfrak{A}_3$ are arbitrary matrices in K.

Consider now the generalized quaternion algebra A over K consisting of the elements $\alpha = a_0\epsilon_0 + a_1\epsilon_1 + a_2\epsilon_2 + a_3\epsilon_3$ with arbitrary a_0, a_1, a_2, a_3 in K, where ϵ_0 is the unit and $\epsilon_1^2 = rs\epsilon_0$, $\epsilon_2^2 = -r\epsilon_0$, $\epsilon_1\epsilon_2 = -\epsilon_2\epsilon_1 = s\epsilon_3$. We denote by $\hat{\alpha} = a_0\epsilon_0 - a_1\epsilon_1 - a_2\epsilon_2 - a_3\epsilon_3$ the conjugate quaternion. There exists the well-known representation of A, of degree 2, defined by

$$\epsilon_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \epsilon_1 = \sqrt{rs}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{74}$$

$$\epsilon_2 = \sqrt{r}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \epsilon_3 = \sqrt{s}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Then obviously

$$\mathfrak{M} = \sum_{k=0}^{3} \mathfrak{A}_k \times \epsilon_k, \tag{75}$$

where $\mathfrak{A}_k \times \epsilon_k$ denotes the Kronecker product of the matrices $\mathfrak{A}_k$ and ϵ_k; and consequently $\mathfrak{C}^{-1}R\mathfrak{C}$ is the ring of all matrices $M = (\alpha_{kl})$ of n rows and columns with arbitrary elements α_{kl} of A.

By (70) and (73), the condition (71) may be expressed in the form

$$\mathfrak{M}\tilde{\mathfrak{M}} = \mathfrak{E} \tag{76}$$

with

$$\tilde{\mathfrak{M}} = \mathfrak{C}^{-1}\tilde{\mathfrak{B}}\mathfrak{C} = \mathfrak{J}^{-1}\mathfrak{M}'\mathfrak{J} = \mathfrak{A}'_0 \times \epsilon_0 - \sum_{k=1}^{3} \mathfrak{A}'_k \times \epsilon_k;$$

hence $\tilde{\mathfrak{M}}$, written as a quaternion matrix, is the transpose $(\hat{\alpha}_{lk})$ of the conjugate $(\hat{\alpha}_{kl})$ of M. By (72) and Section **25**, the group $\Delta_0(\mathfrak{G}, \mathfrak{H})$ consists of all matrices $\mathfrak{M}$ of the form (75), such that (76) is satisfied and $\mathfrak{C}\mathfrak{M}\mathfrak{C}^{-1}$ is integral. On the other hand, the solutions of (76) with integral $\mathfrak{A}_k$ $(k = 0, \cdots, 3)$ in K constitute also a subgroup $\Delta_0(r, s)$ of the homogeneous symplectic group Ω_0. It follows from the argument at the end of Section **26**, that the two groups $\Delta_0(\mathfrak{G}, \mathfrak{H})$ and $\Delta_0(r, s)$ are commensurable. The problem

of the commensurability of two groups $\Delta(\mathfrak{G}, \mathfrak{H})$ and $\Delta(\mathfrak{G}_1, \mathfrak{H}_1)$ is therefore reduced to the corresponding problem for $\Delta_0(r, s)$ and $\Delta_0(r_1, s_1)$. The solution will be given in Chapter IX.

VI. THE FUNDAMENTAL DOMAIN OF THE MODULAR GROUP.

28. We shall construct a fundamental F for any group $\Delta(\mathfrak{G}, \mathfrak{H})$ and, in particular, for the modular group Γ. The application of the general method of Chapter IV would lead to a rather complicated shape of the frontier of F, and it would then be difficult to prove that $\Delta(\mathfrak{G}, \mathfrak{H})$ is a group of the first kind. Therefore we shall use another procedure applying the special arithmetic properties of these groups.

LEMMA 6. *The equation*

$$\mathfrak{S} = \begin{pmatrix} \mathfrak{Y}^{-1} & 0 \\ 0 & \mathfrak{Y} \end{pmatrix} \begin{bmatrix} \mathfrak{E} & -\mathfrak{X} \\ 0 & \mathfrak{E} \end{bmatrix} \tag{77}$$

defines a mapping of the space H of the matrices $\mathfrak{Z} = \mathfrak{X} + i\mathfrak{Y}$ onto the space S of the symplectic positive symmetric matrices $\mathfrak{S}$. Any symplectic transformation $\mathfrak{Z}^* = (\mathfrak{A}\mathfrak{Z} + \mathfrak{B})(\mathfrak{C}\mathfrak{Z} + \mathfrak{D})^{-1}$ with the matrix $\mathfrak{M}$ induces in S the transformation $\mathfrak{S}^* = \mathfrak{S}[\mathfrak{M}^{-1}]$.

Let

$$\mathfrak{S} = \begin{pmatrix} \mathfrak{S}_1 & \mathfrak{S}_{12} \\ \mathfrak{S}'_{12} & \mathfrak{S}_2 \end{pmatrix}$$

be an arbitrary point of S. Since $\mathfrak{S} > 0$, the inequality $\mathfrak{S}_1 > 0$ holds, whence $\mathfrak{S}_1^{-1} = \mathfrak{Y} > 0$. Moreover $\mathfrak{S}_1\mathfrak{S}'_{12} = \mathfrak{S}_{12}\mathfrak{S}_1$ and $\mathfrak{S}_1\mathfrak{S}_2 - \mathfrak{S}_{12}^2 = \mathfrak{E}$, whence $-\mathfrak{S}_1^{-1}\mathfrak{S}_{12} = \mathfrak{X} = \mathfrak{X}'$, $\mathfrak{S}_{12} = -\mathfrak{Y}^{-1}\mathfrak{X}$ and $\mathfrak{S}_2 = \mathfrak{Y} + \mathfrak{Y}^{-1}[\mathfrak{X}]$. This proves the first assertion of the lemma.

The relationship (77) can be written

$$\mathfrak{S}[\mathfrak{w}] = \mathfrak{Y}^{-1}[\mathfrak{u} - \mathfrak{X}\mathfrak{v}] + \mathfrak{Y}[\mathfrak{v}] = \mathfrak{Y}^{-1}\{\mathfrak{u} - \mathfrak{Z}\mathfrak{v}\} \tag{78}$$

with an arbitrary real column

$$\mathfrak{w}^{(2n)} = \begin{pmatrix} \mathfrak{u}^{(n)} \\ \mathfrak{v}^{(n)} \end{pmatrix}.$$

For the symplectic transformation $\mathfrak{Z}^* = \mathfrak{X}^* + i\mathfrak{Y}^* = (\mathfrak{A}\mathfrak{Z} + \mathfrak{B})(\mathfrak{C}\mathfrak{Z} + \mathfrak{D})^{-1}$ we obtain, by (59),

$$\mathfrak{Y}^{*-1} = \mathfrak{Y}^{-1}\{\mathfrak{Z}\mathfrak{C}' + \mathfrak{D}'\}$$

$$\mathfrak{Y}^{*-1}\{\mathfrak{u} - \mathfrak{Z}^*\mathfrak{v}\} = \mathfrak{Y}^{-1}\{(\mathfrak{C}\mathfrak{Z} + \mathfrak{D})'\mathfrak{u} - (\mathfrak{A}\mathfrak{Z} + \mathfrak{B})'\mathfrak{v}\} = \mathfrak{Y}^{-1}\{\mathfrak{u}^* - \mathfrak{Z}\mathfrak{v}^*\},$$

where $\mathfrak{u}^* = \mathfrak{D}'\mathfrak{u} - \mathfrak{B}'\mathfrak{v}$ and $\mathfrak{v}^* = -\mathfrak{C}'\mathfrak{u} + \mathfrak{A}'\mathfrak{v}$. Now the second part of the lemma follows from (8) and (78).

29. Let P be the space of all positive symmetric matrices $\mathfrak{T}^{(m)}$ with real elements. Any $\mathfrak{T}$ in P may be uniquely expressed as $\mathfrak{T} = \mathfrak{P}[\mathfrak{Q}]$, where $\mathfrak{P} = [p_1, \cdots, p_m]$ is a diagonal matrix with $p_k > 0$ $(k = 1, \cdots, m)$ and $\mathfrak{Q} = (q_{kl})$ is a triangular matrix with $q_{kl} = 0$ for $k > l$ and $q_{kk} = 1$. If t is any positive number, the inequalities

$$0 < p_k \leq t p_{k+1}, \qquad -t \leq q_{kl} \leq t \qquad (1 \leq k < l \leq m) \tag{79}$$

define a compact domain $Q(t)$ in P, and any given compact set in P is contained in $Q(t)$ for sufficiently large values of t.

Let U denote the group of all different transformations $\mathfrak{T} \to \mathfrak{T}[\mathfrak{U}]$, where $\mathfrak{U}$ runs over the unimodular matrices, i. e., the matrices with rational integral elements and determinant ± 1. On account of Minkowski's theory of reduction, there exists in P a fundamental domain R with respect to U, defined by a finite number of inequalities

$$L_r(\mathfrak{T}) \geq 0 \qquad (r = 1, 2, \cdots, g), \tag{80}$$

where $L_r(\mathfrak{T})$ denotes a certain homogeneous linear function of the elements of $\mathfrak{T}$ with rational coefficients. A point $\mathfrak{T}$ lies on the frontier of R if, and only if, the conditions (80) are fulfilled with at least one sign of equality. The images of R under U cover the whole space P without gaps and overlappings. Only a finite number of images enter into any compact part of P, and only a finite number of images are neighbors of R.

Most of the results of the theory of reduction are simple consequences of the following two known lemmata.

LEMMA 7. *There exists a positive number τ_1 depending only upon m, such that R is contained in $Q(\tau_1)$.*

LEMMA 8. *Let $\mathfrak{T}_1, \mathfrak{T}_2$ be two points of $Q(t)$ and let $\mathfrak{T}_2 = \mathfrak{T}_1[\mathfrak{F}]$, where $\mathfrak{F}$ is a matrix with rational integral elements f_{kl}. Then*

$$-\tau \leq f_{kl} \leq \tau \qquad (k, l = 1, \cdots, m),$$

where τ is a positive number depending only upon t, m and the determinant $|\mathfrak{F}|$.

Let $t_1, \cdots, t_m$ be the diagonal elements of a point $\mathfrak{T} = \mathfrak{P}[\mathfrak{Q}]$ of R. Then

$$t_l = \sum_{k=1}^{m} p_k q_{kl}^2 = p_l \left(1 + \sum_{k=1}^{l-1} \frac{p_k}{p_l} q_{kl}^2\right) \qquad (l = 1, \cdots, m)$$

and $p_1 p_2 \cdots p_m = |\mathfrak{T}|$. By (79) and Lemma 7, the inequality

$$\prod_{k=1}^{m} t_k < \tau_2 |\mathfrak{T}| \tag{81}$$

follows, where τ_2 depends only upon m.

30. The general method for constructing a fundamental domain with respect to any discontinuous group Δ uses the minimum of the distance $\rho(\mathfrak{Z}^*, \mathfrak{Z}_1)$, where $\mathfrak{Z}_1$ is given and $\mathfrak{Z}^* = (\mathfrak{A}\mathfrak{Z} + \mathfrak{B})(\mathfrak{C}\mathfrak{Z} + \mathfrak{D})^{-1}$ runs over all the images of $\mathfrak{Z}$ under Δ.

Let us now choose in particular $\mathfrak{Z}_1 = i\lambda\mathfrak{E}$, with a positive scalar factor λ; we shall investigate the asymptotic behavior of the distance $\rho(\mathfrak{Z}, \mathfrak{Z}_1)$ for $\lambda \to \infty$. By (38), we have

$$\rho^2(\mathfrak{Z}, \mathfrak{Z}_1) = \sum_{k=1}^{n} \log^2 \frac{1 + r_k^{\frac{1}{2}}}{1 - r_k^{\frac{1}{2}}},$$

where $r_1, \cdots, r_n$ denote the characteristic roots of the cross-ratio

$$\mathfrak{R} = (\mathfrak{Z} - \mathfrak{Z}_1)(\mathfrak{Z} - \bar{\mathfrak{Z}}_1)^{-1}(\bar{\mathfrak{Z}} - \bar{\mathfrak{Z}}_1)(\bar{\mathfrak{Z}} - \mathfrak{Z}_1)^{-1} = \mathfrak{E} + 2i\lambda^{-1}(\mathfrak{Z} - \bar{\mathfrak{Z}}) + \cdots.$$

If $s_1, \cdots, s_n$ are the characteristic roots of the imaginary part $\mathfrak{Y}$ of $\mathfrak{Z}$, we obtain

$$r_k = 1 - 4s_k\lambda^{-1} + \cdots \qquad (k = 1, \cdots, n),$$

whence

$$\rho^2(\mathfrak{Z}, i\lambda\mathfrak{E}) = \sum_{k=1}^{n} \log^2 (s_k^{-1}\lambda) + \omega(\lambda),$$

where $\omega(\lambda)$ is a power series in λ^{-1} without constant term. Consequently

$$\lim_{\lambda\to\infty} (\rho(\mathfrak{Z}, i\lambda\mathfrak{E}) - n^{\frac{1}{2}} \log \lambda) = n^{-\frac{1}{2}} \log |\mathfrak{Y}|^{-1}.$$

This suggests a consideration of the minimum of $|\mathfrak{Y}^*|^{-1}$.

Denoting by the sign abs $\mathfrak{R}$ the absolute value of the determinant of a matrix $\mathfrak{R}$, we have, by (59),

$$|\mathfrak{Y}^*|^{-1} = |\mathfrak{Y}|^{-1} \operatorname{abs} (\mathfrak{C}\mathfrak{Z} + \mathfrak{D})^2.$$

In order to obtain the minimum of $|\mathfrak{Y}^*|^{-1}$, we have therefore to determine the minimum of abs $(\mathfrak{C}\mathfrak{Z} + \mathfrak{D})$. The existence of this minimum is by no means trivial; we shall prove it now in the case of the modular group Γ.

Let

$$\mathfrak{M} = \begin{pmatrix} \mathfrak{A} & \mathfrak{B} \\ \mathfrak{C} & \mathfrak{D} \end{pmatrix}, \qquad \mathfrak{M}_0 = \begin{pmatrix} \mathfrak{A}_0 & \mathfrak{B}_0 \\ \mathfrak{C}_0 & \mathfrak{D}_0 \end{pmatrix}$$

be two elements of the homogeneous modular group Γ_0 and let

$$\mathfrak{M}_0\mathfrak{M}^{-1} = \mathfrak{M}_1 = \begin{pmatrix} \mathfrak{A}_1 & \mathfrak{B}_1 \\ \mathfrak{C}_1 & \mathfrak{D}_1 \end{pmatrix}.$$

We assume that the equation

$$\text{abs}\,(\mathfrak{C}\mathfrak{Z} + \mathfrak{D}) = \text{abs}\,(\mathfrak{C}_0\mathfrak{Z} + \mathfrak{D}_0) \tag{82}$$

holds identically for all $\mathfrak{Z}$ in H. Introducing $(\mathfrak{D}'\mathfrak{Z} - \mathfrak{B}')(-\mathfrak{C}'\mathfrak{Z} + \mathfrak{A}')^{-1}$ instead of $\mathfrak{Z}$, we obtain the necessary and sufficient condition

$$\text{abs}\,(\mathfrak{C}_1\mathfrak{Z} + \mathfrak{D}_1) = 1. \tag{83}$$

Since $|\,\mathfrak{C}_1\mathfrak{Z} + \mathfrak{D}_1\,|$ is an analytic function of the elements z_{kl} of $\mathfrak{Z}$, we infer that $|\,\mathfrak{C}_1\mathfrak{Z} + \mathfrak{D}_1\,| = c$, identically, for all complex symmetric matrices $\mathfrak{Z}$, with a constant c of absolute value 1. Putting $\mathfrak{Z} = 0$, we find $|\,\mathfrak{D}_1\,| = c$. On the other hand, the elements of $\mathfrak{D}_1$ are rational integers; consequently $\mathfrak{D}_1$ is a unimodular matrix $\mathfrak{U}$ and $c = \pm 1$. Calculating the linear terms in the identity $|\,\mathfrak{D}_1^{-1}\mathfrak{C}_1\mathfrak{Z} + \mathfrak{E}\,| = 1$, we obtain $\sigma(\mathfrak{D}_1^{-1}\mathfrak{C}_1\mathfrak{Z}) = 0$. But the matrix $\mathfrak{D}_1^{-1}\mathfrak{C}_1$ is symmetric and therefore $\mathfrak{C}_1 = 0$.

Let now $\mathfrak{M}_1$ be any modular matrix with $\mathfrak{C}_1 = 0$. The general form is

$$\mathfrak{M}_1 = \begin{pmatrix} \mathfrak{U}'^{-1} & \mathfrak{B}\mathfrak{U} \\ 0 & \mathfrak{U} \end{pmatrix} \tag{84}$$

with unimodular $\mathfrak{U}$ and integral symmetric $\mathfrak{B}$. Obviously (79) is satisfied, and $\mathfrak{M}_0 = \mathfrak{M}_1\mathfrak{M}$, with an arbitrary modular matrix $\mathfrak{M}$, gives the general solution of (82). Then

$$\mathfrak{C}_0 = \mathfrak{U}\mathfrak{C}, \qquad \mathfrak{D}_0 = \mathfrak{U}\mathfrak{D}. \tag{85}$$

It is also easily seen that $\mathfrak{M}_1 = \mathfrak{M}_0\mathfrak{M}^{-1}$ has always the form (84), if the second matrix rows $(\mathfrak{C}\mathfrak{D})$ and $(\mathfrak{C}_0\mathfrak{D}_0)$ of two modular matrices $\mathfrak{M}$ and $\mathfrak{M}_0$ are connected by (85), with unimodular $\mathfrak{U}$. The two pairs $\mathfrak{C}, \mathfrak{D}$ and $\mathfrak{C}_0, \mathfrak{D}_0$ are called *associate*.

Denoting by $\mathfrak{Y}^*$ and $\mathfrak{Y}^*_0$ the imaginary parts of

$$\mathfrak{Z}^* = (\mathfrak{A}\mathfrak{Z} + \mathfrak{B})(\mathfrak{C}\mathfrak{Z} + \mathfrak{D})^{-1} \quad \text{and} \quad \mathfrak{Z}^*_0 = (\mathfrak{A}_0\mathfrak{Z} + \mathfrak{B}_0)(\mathfrak{C}_0\mathfrak{Z} + \mathfrak{D}_0)^{-1},$$

we obtain, by (59),

$$\mathfrak{Y}^{-1}\{\mathfrak{Z}\mathfrak{C}'_0 + \mathfrak{D}'_0\} = \mathfrak{Y}^{*-1}[\mathfrak{U}'] = \mathfrak{Y}^{*}{}_0^{-1}. \tag{86}$$

Let $\mathfrak{Z}$ be a given point of H. For any modular matrix $\mathfrak{M}$, we choose a unimodular matrix $\mathfrak{U}$, such that $\mathfrak{Y}^{*-1}[\mathfrak{U}'] = \mathfrak{Y}^{*}{}_0^{-1}$ lies in the Minkowski domain R of Section **29**. We shall now prove that abs $(\mathfrak{C}_0\mathfrak{Z} + \mathfrak{D}_0)$ tends to infinity, if $(\mathfrak{C}_0\mathfrak{D}_0)$ runs over all second matrix rows of modular matrices with the required property of $\mathfrak{Y}^*_0$.

Let $y_1, \cdots, y_n$ denote the diagonal elements of $\mathfrak{Y}^{*}{}_0^{-1}$ and $\mathfrak{c}_l, \mathfrak{d}_l$ the l-th columns of $\mathfrak{C}'_0, \mathfrak{D}'_0$ $(l = 1, \cdots, n)$. By (86),

$$\begin{aligned} \mathfrak{Y}^{*}{}_0^{-1} &= \mathfrak{Y}^{-1}[\mathfrak{X}\mathfrak{C}'_0 + \mathfrak{D}'_0] + \mathfrak{Y}[\mathfrak{C}'_0]; \\ y_l &= \mathfrak{Y}^{-1}[\mathfrak{X}\mathfrak{c}_l + \mathfrak{d}_l] + \mathfrak{Y}[\mathfrak{c}_l] \qquad (l = 1, \cdots, n). \end{aligned} \tag{87}$$

Consider all solutions $\mathfrak{C}_0, \mathfrak{D}_0$ of the inequality $\operatorname{abs}(\mathfrak{C}_0\mathfrak{Z} + \mathfrak{D}_0)^2 < a$, where a is an arbitrary positive constant. By (81) and (86),

$$\prod_{l=1}^{n} y_l < \tau_2 a \,|\, \mathfrak{Y} \,|^{-1};$$

by (87),

$$\mathfrak{Y}[\mathfrak{c}_l] \leq y_l, \qquad \mathfrak{Y}^{-1}[\mathfrak{X}\mathfrak{c}_l + \mathfrak{d}_l] \leq y_l \qquad (l = 1, \cdots, n), \tag{88}$$

and $\mathfrak{Y}^{-1}[\mathfrak{d}_l] = y_l$ in the case $\mathfrak{c}_l = 0$. Since $|\, \mathfrak{C}_0\mathfrak{Z} + \mathfrak{D}_0 \,| \neq 0$, the columns $\mathfrak{c}_l, \mathfrak{d}_l$ are not both 0; hence $y_1 > y$, where y is a positive number depending only upon $\mathfrak{Y}$, and

$$y_l < \tau_2 a y^{1-n} \,|\, \mathfrak{Y} \,|^{-1}.$$

By (88), we obtain only a finite number of pairs $\mathfrak{C}_0, \mathfrak{D}_0$.

On account of (82), the existence of a modular transformation $\mathfrak{Z}^* = (\mathfrak{A}\mathfrak{Z} + \mathfrak{B})(\mathfrak{C}\mathfrak{Z} + \mathfrak{D})^{-1}$ with the minimum value of $\operatorname{abs}(\mathfrak{C}\mathfrak{Z} + \mathfrak{D})$ is established for any $\mathfrak{Z}$ in H. We determine again $\mathfrak{U}$ by the condition that $\mathfrak{Y}^{*-1}[\mathfrak{U}']$ is a point of R and define $\mathfrak{Z}^*{}_0 = (\mathfrak{A}_0\mathfrak{Z} + \mathfrak{B}_0)(\mathfrak{C}_0\mathfrak{Z} + \mathfrak{D}_0)^{-1} = \mathfrak{Z}^*[\mathfrak{U}^{-1}] + \mathfrak{B}$, by (84), where $\mathfrak{B}$ is an arbitrary symmetric matrix with rational integral elements. We may choose $\mathfrak{B}$ such that all elements of the real part of $\mathfrak{Z}^*{}_0$ lie in the interval $-\frac{1}{2} \leq x \leq \frac{1}{2}$.

31. Let F be the set of all points $\mathfrak{Z} = \mathfrak{X} + i\mathfrak{Y}$ of H satisfying the following three conditions:

$$\operatorname{abs}(\mathfrak{C}\mathfrak{Z} + \mathfrak{D}) \geq 1 \tag{89}$$

for all modular transformations $\mathfrak{Z}^* = (\mathfrak{A}\mathfrak{Z} + \mathfrak{B})(\mathfrak{C}\mathfrak{Z} + \mathfrak{D})^{-1}$;

$$L_r(\mathfrak{Y}^{-1}) \geq 0 \qquad (r = 1, \cdots, g); \tag{90}$$

$$x_{kl} \geq -\tfrac{1}{2}, \qquad -x_{kl} \geq -\tfrac{1}{2} \qquad (1 \leq k \leq l \leq n) \tag{91}$$

for the elements x_{kl} of $\mathfrak{X}$. In (89), we shall omit the trivial case $\mathfrak{C} = 0$, since the corresponding condition $\operatorname{abs}(\mathfrak{D}) \geq 1$ holds identically for all $\mathfrak{Z}$. By the result of the preceding section, the images of F under Γ cover the whole space H.

We write $\mathfrak{Y}^{-1} = \mathfrak{P}[\mathfrak{Q}]$ with $\mathfrak{P} = [p_1, \cdots, p_n]$, $\mathfrak{Q} = (q_{kl})$, $q_{kl} = 0$ $(1 \leq l < k \leq n)$, $q_{kk} = 1$ $(k = 1, \cdots, n)$ and define $\mathfrak{W}^{(n)} = (w_{kl})$ with $w_{kl} = 0$ $(k + l \neq n + 1)$, $w_{kl} = 1$ $(k + l = n + 1)$,

$$\mathfrak{P}_1 = \begin{pmatrix} \mathfrak{P} & 0 \\ 0 & \mathfrak{P}^{-1}[\mathfrak{W}] \end{pmatrix}, \qquad \mathfrak{Q}_1 = \begin{pmatrix} \mathfrak{Q} & -\mathfrak{Q}\mathfrak{X} \\ 0 & \mathfrak{W}\mathfrak{Q}'^{-1}\mathfrak{W} \end{pmatrix}, \qquad \mathfrak{W}_1 = \begin{pmatrix} \mathfrak{E} & 0 \\ 0 & \mathfrak{W} \end{pmatrix}.$$

By (77),

$$\mathfrak{S}[\mathfrak{W}_1] = \mathfrak{P}_1[\mathfrak{Q}_1]. \tag{92}$$

By Lemma 7 and (91), the absolute values of the elements of the triangular matrix $\mathfrak{Q}_1$ are less than a number depending only upon n. Denoting the diagonal elements of $\mathfrak{P}_1$ by $d_1, \cdots, d_{2n}$, we have

$$d_k d_{k+1}^{-1} = p_k p_{k+1}^{-1} \ (1 \leq k < n), \ = p_n^2 \ (k = n), \ = p_{2n-k} p_{2n-k+1}^{-1} \ (n < k < 2n).$$

By Lemma 7,

$$0 < d_k \leq \tau_1 d_{k+1} \qquad (k \neq n), \tag{93}$$

where τ_1 depends only upon $m = 2n$.

We apply (89) for the particular modular transformation with

$$\mathfrak{A} = \begin{pmatrix} \mathfrak{E}^{(n-1)} & 0 \\ 0 & 0 \end{pmatrix}, \quad \mathfrak{B} = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}, \quad \mathfrak{C} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathfrak{D} = \begin{pmatrix} \mathfrak{E}^{(n-1)} & 0 \\ 0 & 0 \end{pmatrix}.$$

Denoting by $x_n + iy_n$ the last diagonal element of $\mathfrak{Z}$, we obtain the inequality $x_n^2 + y_n^2 \geq 1$. By (91), we have moreover $x_n^2 \leq \frac{1}{4}$, whence $y_n^2 \geq \frac{3}{4}$. But $\mathfrak{Y} = \mathfrak{P}^{-1}[\mathfrak{Q}'^{-1}]$, and consequently $y_n = p_n^{-1}$,

$$0 < d_n \leq (4/3)\, d_{n+1}. \tag{94}$$

By (92), (93) and (94), $\mathfrak{S}[\mathfrak{W}_1]$ is contained in a domain $Q(\tau_3)$ of the space P of positive symmetric matrices with $2n$ rows, where τ_3 depends only upon n.

Consider now any modular transformation $\mathfrak{Z}^* = (\mathfrak{A}\mathfrak{Z} + \mathfrak{B})(\mathfrak{C}\mathfrak{Z} + \mathfrak{D})^{-1}$ different from identity, i. e.,

$$\mathfrak{M} = \begin{pmatrix} \mathfrak{A} & \mathfrak{B} \\ \mathfrak{C} & \mathfrak{D} \end{pmatrix} \neq \pm \mathfrak{E},$$

and assume that $\mathfrak{Z}$ and $\mathfrak{Z}^*$ both are points of F. By Lemma 6, $\mathfrak{S} = \mathfrak{S}^*[\mathfrak{M}]$,

$$\mathfrak{S}[\mathfrak{W}_1] = \mathfrak{S}^*[\mathfrak{W}_1][\mathfrak{W}_1\mathfrak{M}\mathfrak{W}_1].$$

Applying Lemma 8 with

$$\mathfrak{T}_1 = \mathfrak{S}^*[\mathfrak{W}_1], \quad \mathfrak{T}_2 = \mathfrak{S}[\mathfrak{W}_1], \quad \mathfrak{F} = \mathfrak{W}_1\mathfrak{M}\mathfrak{W}_1, \quad |\mathfrak{F}| = 1,$$

we conclude that $\mathfrak{M}$ belongs to a finite set of modular matrices $\mathfrak{M}_1, \cdots, \mathfrak{M}_h$, independent of $\mathfrak{Z}$ and $\mathfrak{Z}^*$.

On account of the minimum property of $|\mathfrak{Y}|^{-1}$, we have $|\mathfrak{Y}| = |\mathfrak{Y}^*|$,

whence abs $(\mathfrak{C}\mathfrak{Z}+\mathfrak{D})=1$. If $\mathfrak{C}\neq 0$, then the sign of equality is true in one of the conditions (89). If $\mathfrak{C}=0$, then $\mathfrak{M}$ has the form (84),

$$\mathfrak{M}=\begin{pmatrix}\mathfrak{U}'^{-1} & \mathfrak{B}\mathfrak{U}\\ 0 & \mathfrak{U}\end{pmatrix}$$

and $\mathfrak{Y}^{*-1}=\mathfrak{Y}^{-1}[\mathfrak{U}']$. In the case $\mathfrak{U}\neq\pm\mathfrak{E}$, the sign of equality is true in one of the conditions (90). In the case $\mathfrak{U}=\pm\mathfrak{E}$, we have $\mathfrak{Z}^*=\mathfrak{Z}+\mathfrak{B}$, $\mathfrak{X}^*=\mathfrak{X}+\mathfrak{B}$ with integral symmetric $\mathfrak{B}\neq 0$, and then the sign of equality holds in one of the conditions (91).

We have proved the following statement: If two points $\mathfrak{Z}$ and $\mathfrak{Z}^*$ of F are equivalent under a modular transformation with the matrix $\mathfrak{M}\neq\pm\mathfrak{E}$, then $\mathfrak{M}$ is one of the matrices

$$\mathfrak{M}_s=\begin{pmatrix}\mathfrak{A}_s & \mathfrak{B}_s\\ \mathfrak{C}_s & \mathfrak{D}_s\end{pmatrix} \qquad (s=1,\cdots,h)$$

and the conditions

$$\text{(95)} \qquad \text{abs}\,(\mathfrak{C}_s\mathfrak{Z}+\mathfrak{D}_s)\geq 1 \qquad (s=1,\cdots,h),$$

$$\text{(96)} \qquad L_r(\mathfrak{Y}^{-1})\geq 0 \qquad (r=1,\cdots,g),$$

$$\text{(97)} \qquad x_{kl}\geq -\tfrac{1}{2},\quad -x_{kl}\geq -\tfrac{1}{2} \qquad (1\leq k\leq l\leq n)$$

are fulfilled with at least one sign of equality.

32. Since (95) is contained in (89), every point $\mathfrak{Z}$ of F satisfies (95), (96) and (97). We prove, now, that the converse is true, namely, that all the inequalities (89) follow from (95), (96) and (97). We shall demonstrate at the same time, that then the stronger inequalities abs $(\mathfrak{C}\mathfrak{Z}+\mathfrak{D})>1$ hold, if $\mathfrak{C},\mathfrak{D}$ is not associate with one of the pairs $\mathfrak{C}_s,\mathfrak{D}_s$ $(s=1,\cdots,h)$.

LEMMA 9. *Let $\mathfrak{Z}=\mathfrak{X}+i\mathfrak{Y}$ be a point of H and $\mathfrak{Z}_\lambda=\mathfrak{X}+i\lambda\mathfrak{Y}$, with an arbitrary scalar factor λ. If $(\mathfrak{C}\mathfrak{D})$ is the second matrix row of any symplectic matrix and $\mathfrak{C}\neq 0$, then the inequality* abs $(\mathfrak{C}\mathfrak{Z}_\lambda+\mathfrak{D})>$ abs $(\mathfrak{C}\mathfrak{Z}_\mu+\mathfrak{D})$ *holds for $\lambda>\mu>0$.*

The determinant $|\,\mathfrak{C}\mathfrak{Z}^\lambda+\mathfrak{D}\,|=\phi(\lambda)$ is a polynomial in λ. For any λ with positive real part, $\mathfrak{Z}_\lambda$ is a point of H and consequently $\phi(\lambda)\neq 0$. Moreover $\overline{\phi(\lambda)}=\phi(-\bar{\lambda})$; hence all zeros of $\phi(\lambda)$ are pure imaginary. This proves that for real λ the expression abs $(\mathfrak{C}\mathfrak{Z}_\lambda+\mathfrak{D})^2$ is a polynomial in λ^2 with real non-negative coefficients. It remains only to prove that this polynomial is not identically constant.

Assume now that

$$\text{abs}\,(\mathfrak{C}\mathfrak{Z}^\lambda+\mathfrak{D})=\text{abs}\,(\mathfrak{C}\mathfrak{Z}_0+\mathfrak{D})=\text{abs}\,(\mathfrak{C}\mathfrak{X}+\mathfrak{D}),$$

identically in λ. Then abs $(\mathfrak{C}\mathfrak{X}+\mathfrak{D}) \neq 0$ and

$$\text{abs}\,(\mathfrak{Y}^{-1}+i\lambda(\mathfrak{C}\mathfrak{X}+\mathfrak{D})^{-1}\mathfrak{C}) = |\,\mathfrak{Y}\,|^{-1}. \tag{98}$$

Since $(\mathfrak{C}\mathfrak{X}+\mathfrak{D})^{-1}\mathfrak{C}$ is a real symmetric matrix and $\mathfrak{Y}^{-1}>0$, there exists a real matrix $\mathfrak{R}^{(n)}$ and a diagonal matrix $\mathfrak{N}=[r_1,\cdots,r_n]$, such that

$$\mathfrak{Y}^{-1}[\mathfrak{R}]=\mathfrak{E}, \qquad ((\mathfrak{C}\mathfrak{X}+\mathfrak{D})^{-1}\mathfrak{C})[\mathfrak{R}]=\mathfrak{N}. \tag{99}$$

By (98) and (99),

$$\prod_{k=1}^{n}(1+r_k^2\lambda^2)=1$$

for all real values of λ, and consequently $\mathfrak{N}=0$, $\mathfrak{C}=0$, which is a contradiction. This completes the proof of the lemma.

We denote again by $(\mathfrak{C}\mathfrak{D})$ the second matrix rows of the modular matrices. Let $\mathfrak{Z}$ be a point of H satisfying all conditions (95), (96) and (97), and assume that the inequality abs $(\mathfrak{C}\mathfrak{Z}+\mathfrak{D})\leq 1$ holds for at least one pair $\mathfrak{C},\mathfrak{D}$, where $\mathfrak{C}\neq 0$ and $\mathfrak{C},\mathfrak{D}$ is not associate with one of the pairs $\mathfrak{C}_s,\mathfrak{D}_s$ $(s=1,\cdots,h)$. By the result of Section **30**, only a finite number of non-associate pairs $\mathfrak{C},\mathfrak{D}$ fulfill that inequality. By Lemma 9, there exists a number $\lambda\geq 1$, such that abs $(\mathfrak{C}\mathfrak{Z}_\lambda+\mathfrak{D})\geq 1$ for all $\mathfrak{C},\mathfrak{D}$ and abs $(\mathfrak{C}\mathfrak{Z}_\lambda+\mathfrak{D})=1$ for $\mathfrak{C}=\mathfrak{C}_0$, $\mathfrak{D}=\mathfrak{D}_0$, where $\mathfrak{C}_0\neq 0$ and $\mathfrak{C}_0,\mathfrak{D}_0$ is not associate with one of the pairs $\mathfrak{C}_s,\mathfrak{D}_s$. Since $\mathfrak{Z}_\lambda$ has the real part $\mathfrak{X}$ and $L_r(\lambda^{-1}\mathfrak{Y}^{-1})=\lambda^{-1}L_r(\mathfrak{Y}^{-1})$, all conditions (89), (90) and (91) are satisfied for $\mathfrak{Z}_\lambda$ instead of $\mathfrak{Z}$; hence $\mathfrak{Z}_\lambda$ is a point of F. On the other hand, the expression abs $(\mathfrak{C}\mathfrak{Z}_\lambda+\mathfrak{D})$ attains its minimum 1 for $\mathfrak{C}=\mathfrak{C}_0$, $\mathfrak{D}=\mathfrak{D}_0$, and consequently there exists a modular transformation $\mathfrak{Z}^*_\lambda=(\mathfrak{A}\mathfrak{Z}_\lambda+\mathfrak{B})(\mathfrak{C}\mathfrak{Z}_\lambda+\mathfrak{D})^{-1}$, such that $\mathfrak{C},\mathfrak{D}$ is associate with $\mathfrak{C}_0,\mathfrak{D}_0$ and $\mathfrak{Z}^*_\lambda$ is a point of F. By Section **31**, $\mathfrak{C}=\mathfrak{C}_s$, $\mathfrak{D}=\mathfrak{D}_s$, and this is a contradiction. Consequently abs $(\mathfrak{C}\mathfrak{Z}+\mathfrak{D})\geq 1$ for all $\mathfrak{C},\mathfrak{D}$ and abs $(\mathfrak{C}\mathfrak{Z}+\mathfrak{D})>1$, if $\mathfrak{C}\neq 0$ and $\mathfrak{C},\mathfrak{D}$ is not associate with one of the pairs $\mathfrak{C}_s,\mathfrak{D}_s$.

33. By the result of the preceding section, F may be defined by the inequalities (95), (96) and (97), in finite number. Obviously F is closed relative to H. It follows from Lemma 9 and the linearity of the conditions (96) and (97), that $\mathfrak{Z}$ is a frontier point of F, if, and only if, (95), (96) and (97) are fulfilled with at least one sign of equality.

Let $F_{\mathfrak{M}}$ be the image of F under the modular transformation $\mathfrak{Z}^*=(\mathfrak{A}\mathfrak{Z}+\mathfrak{B})(\mathfrak{C}\mathfrak{Z}+\mathfrak{D})^{-1}$ with the matrix $\mathfrak{M}\neq\pm\mathfrak{E}$. If F and $F_{\mathfrak{M}}$ have a point $\mathfrak{Z}^*$ in common, then, by Section **31**, $\mathfrak{M}$ is one of the matrices $\mathfrak{M}_1,\cdots,\mathfrak{M}_h$, and $\mathfrak{Z}^*$ is a frontier point of F. Consequently the images $F_{\mathfrak{M}}$ cover H without overlappings, and F has only a finite number of neighbors.

Consider any compact domain G in H, and let G_1 be the corresponding domain in the space S of the matrices $\mathfrak{S}$ defined in Lemma 6. There exists a number t depending only upon G, such that G_1 is contained in the domain $Q(t)$ of Section **29**. We may choose $t \geq \tau_3$, where τ_3 was defined in Section **31**. Let $\mathfrak{Z}^* = (\mathfrak{A}\mathfrak{Z} + \mathfrak{B})(\mathfrak{C}\mathfrak{Z} + \mathfrak{D})^{-1}$ be a common point of $F_{\mathfrak{M}}$ and G. Then $\mathfrak{Z}$ is a point of F and the relationship $\mathfrak{S}^* = \mathfrak{S}[\mathfrak{M}^{-1}]$ holds for the corresponding points $\mathfrak{S}^*$ and $\mathfrak{S}$ of S, by Lemma 6. It follows from the result of Section **31**, that the point $\mathfrak{S}[\mathfrak{W}_1] = \mathfrak{S}^*[\mathfrak{M}\mathfrak{W}_1]$ lies in $Q(t)$. But also $\mathfrak{S}^*$ itself is a point of $Q(t)$, and consequently, by Lemma 8, the matrix $\mathfrak{M}$ belongs to a finite set. This proves that only a finite number of images $F_{\mathfrak{M}}$ enter into the compact domain G.

For the particular value $\mathfrak{Z} = i\mathfrak{E}$, we have $|\mathfrak{C}\mathfrak{Z} + \mathfrak{D}| = c + id$, with rational integers c, d not both 0. Consequently (89) is satisfied. Also (91) holds, since $\mathfrak{X} = 0$. Moreover $\mathfrak{Y}^{-1} = \mathfrak{E}$ is a point of the Minkowski domain R. Consequently $\mathfrak{Z} = i\mathfrak{E}$ is a point of F. By Lemma 9, the whole curve $\mathfrak{Z} = i\lambda\mathfrak{E}$ $(\lambda \geq 1)$ belongs to F. Since λ may be arbitrarily large, the fundamental domain F is not compact. Let G be any compact domain in H and consider the finite set of modular matrices $\mathfrak{M}$, such that $F_{\mathfrak{M}}$ enters into G. The set of images of G under the inverse mappings with the matrices $\mathfrak{M}^{-1}$ constitutes again a compact domain G_0. For sufficiently large values of λ, the point $i\lambda\mathfrak{E}$ of F does not lie in G_0; hence no image of this point lies in G. This proves that the space H is not compact relative to the modular group Γ.

By the results of Section **31**, the matrices $\mathfrak{X}$ and $\mathfrak{Y}^{-1}$ are bounded for all $\mathfrak{Z}$ in F. On account of (57), the integral $V(\Gamma)$ converges.

Theorem 8 is now completely proved.

VII. THE FUNDAMENTAL DOMAIN OF THE GROUP $\Delta(\mathfrak{G}, \mathfrak{H})$.

34. Let K_0 be an algebraic number field, of finite degree g over the field of rational numbers. Let g_1 of the conjugate fields be real and $2g_2$ imaginary, $g = g_1 + 2g_2$. We denote the real conjugate fields by $K_0^{(\alpha)}$ $(\alpha = 1, \cdots, g_1)$ and the pairs of conjugate complex conjugate fields by $K_0^{(\alpha)}$ and $K_0^{(\alpha+g_2)}$ $(\alpha = g_1 + 1, \cdots, g_1 + g_2)$. We consider g_1 positive symmetric matrices $\mathfrak{T}_\alpha$ $(\alpha = 1, \cdots, g_1)$ with real elements and g_2 positive hermitian matrices $\mathfrak{T}_\alpha$ $(\alpha = g_1 + 1, \cdots, g_1 + g_2)$ with complex elements, all of m rows. We denote the systems of $g_1 + g_2$ matrices $\mathfrak{T}_\alpha$ $(\alpha = 1, \cdots, g_1 + g_2)$ more shortly by $\mathfrak{T}$; they form the points of a space P of $\frac{1}{2}g_1 m(m+1) + g_2 m^2$ dimensions.

We have a unique decomposition $\mathfrak{T}_\alpha = \mathfrak{P}_\alpha\{\mathfrak{Q}_\alpha\}$ with a diagonal matrix

$\mathfrak{P}_\alpha = [p_1^{(\alpha)}, \cdots, p_m^{(\alpha)}]$, $p_k^{(\alpha)} > 0$ $(k = 1, \cdots, m)$, and a triangular matrix $\mathfrak{Q}_\alpha = (q_{kl}^{(\alpha)})$, $q_{kl}^{(\alpha)} = 0$ $(k > l)$, $q_{kk}^{(\alpha)} = 1$, where $\mathfrak{Q}_\alpha$ is real for $\alpha = 1, \cdots, g_1$. If t is any positive number > 1, the inequalities

$$0 < p_k^{(\alpha)} \leq t p_{k+1}^{(\alpha)} \quad (k < m), \qquad p_k^{(\alpha)} \leq t p_k^{(\beta)} \quad (k \leq m)$$
$$\operatorname{abs} q_{kl}^{(\alpha)} \leq t \qquad (k < l)$$

with $\alpha, \beta = 1, \cdots, g_1 + g_2$ define a compact domain $Q(t)$ in P.

A matrix $\mathfrak{U}$ with integral elements in K_0 is called unimodular if the determinant $|\mathfrak{U}|$ is an algebraic unit. The unimodular matrices $\mathfrak{U}^{(m)}$ constitute the unimodular group $\mathfrak{U}$ in K_0, of degree m. The center C of U consists of the matrices $\mathfrak{U} = u\mathfrak{E}$, where u is any root of unity in K_0. We denote by U_0 the factor group U/C. Let $\mathfrak{U}_\alpha$ be the conjugate of $\mathfrak{U}$ in $K_0^{(\alpha)}$. The transformation $\mathfrak{T}_\alpha \to \mathfrak{T}_\alpha\{\mathfrak{U}_\alpha\}$ $(\alpha = 1, \cdots, g_1 + g_2)$, or more shortly $\mathfrak{T} \to \mathfrak{T}\{\mathfrak{U}\}$, maps the space P onto itself. This mapping is the identical one, if and only if $\mathfrak{U}$ is an element of C; consequently the transformations $\mathfrak{T} \to \mathfrak{T}\{\mathfrak{U}\}$ give a faithful representation of U_0.

Minkowski's theory of reduction of positive quadratic forms is concerned with the case $g = 1$, the field of rational numbers. The generalization to the case of an arbitrary field K_0 is due to P. Humbert. He obtained the following results:

There exists in P a fundamental domain R with respect to U_0, which is the union of a finite number of convex pyramids. The faces of these pyramids have equations of the form

$$\sum_{\alpha=1}^{g_1+g_2} (\mathfrak{a}'_\alpha \mathfrak{T}_\alpha \bar{\mathfrak{b}}_\alpha + \bar{\mathfrak{a}}'_\alpha \bar{\mathfrak{T}}_\alpha \mathfrak{b}_\alpha) = 0, \tag{100}$$

where $\mathfrak{a}_\alpha$ and $\mathfrak{b}_\alpha$ are the conjugates of two columns $\mathfrak{a} \neq 0$ and $\mathfrak{b} \neq 0$ in K_0; moreover $\mathfrak{a} \neq \lambda\mathfrak{b}$ for every pure imaginary scalar factor λ. Any compact domain in P is covered by a finite number of images $R_\mathfrak{U}$ of R, and R has only a finite number of neighbors $R_\mathfrak{U}$.

LEMMA 10. *There exists a finite set L of matrices $\mathfrak{L}$ with integral elements in K_0 and a positive number τ_4 depending only upon K_0 and m, such that for every $\mathfrak{T}$ in R the point $\mathfrak{T}\{\mathfrak{L}\}$ belongs to $Q(\tau_4)$, with at least one $\mathfrak{L}$ of the set L.*

LEMMA 11. *Let $\mathfrak{T}$ and $\mathfrak{T}\{\mathfrak{F}\}$ be two points of $Q(t)$, where $\mathfrak{F}$ is an integral matrix in K_0, and let ν be the norm of $|\mathfrak{F}|$. Then $\mathfrak{F}$ belongs to a finite set of matrices depending only upon K_0, m, t and ν.*

These statements are generalizations of the lemmata 7 and 8.

35. We assume now that $K_0 = K(\sqrt{-r})$, where K is a totally real algebraic number field of degree h and r a totally positive number of K; then $g_1 = 0$, $g_2 = h$. Let $K^{(\alpha)}$ $(\alpha = 1, \cdots, h)$ be the conjugates of $K = K^{(1)}$, r_α the conjugate of r in $K^{(\alpha)}$ and $K_0{}^{(\alpha)} = K^{(\alpha)}(\sqrt{-r_\alpha})$. In the notation of Chapter V, $\mathfrak{H}$ is a hermitian matrix and $\mathfrak{G}$ is a skew-symmetric matrix, both in K_0, satisfying the condition (67). Let $\mathfrak{H}_\alpha, \mathfrak{G}_\alpha$ be the conjugates of $\mathfrak{H}, \mathfrak{G}$ in $K_0{}^{(\alpha)}$ and s_α the conjugate of s in $K^{(\alpha)}$. We assumed moreover $s > 0$, $\mathfrak{H}_\alpha > 0$ for $\alpha = 2, \cdots, h$; then $s_\alpha < 0$ $(\alpha = 2, \cdots, h)$. The group $\Lambda = \Lambda(\mathfrak{G}, \mathfrak{H})$ consists of all unimodular matrices $\mathfrak{W}$ in K_0 which satisfy $\mathfrak{G}[\mathfrak{W}] = \mathfrak{G}$ and $\mathfrak{H}\{\mathfrak{W}\} = \mathfrak{H}$. We have $\Delta_0 = \mathfrak{C}^{-1}\Lambda\mathfrak{C}$, where $\mathfrak{C}$ is a complex matrix with $\mathfrak{G}[\mathfrak{C}] = \mathfrak{J}$ and $\mathfrak{H}\{\mathfrak{C}\} = is^{\frac{1}{2}}\mathfrak{J}$. Identifying $\mathfrak{W}$ and $-\mathfrak{W}$, we obtain the factor group $\Delta = \Delta(\mathfrak{G}, \mathfrak{H})$ of Δ_0. Obviously this group is not changed, if we replace $\mathfrak{G}, \mathfrak{H}$ by $a\mathfrak{G}, b\mathfrak{H}$ with arbitrary positive rational numbers a, b; consequently we may assume that $\mathfrak{G}$ and $\mathfrak{H}$ have integral elements.

The matrix $\mathfrak{S}$ of (77) is the general solution of $\mathfrak{S}' = \bar{\mathfrak{S}} > 0$, $\mathfrak{J}[\mathfrak{S}] = \mathfrak{J}$, $\mathfrak{J}\{\mathfrak{S}\} = \mathfrak{J}$. Consequently

$$\mathfrak{T}_1 = \mathfrak{S}\{\mathfrak{C}^{-1}\} \tag{101}$$

is the general solution of

$$\mathfrak{T}'_1 = \bar{\mathfrak{T}}_1 > 0, \quad \mathfrak{T}_1\bar{\mathfrak{G}}^{-1}\bar{\mathfrak{T}}_1 = -\mathfrak{G}, \quad \mathfrak{T}_1\mathfrak{H}^{-1}\mathfrak{T}_1 = s^{-1}\mathfrak{H}. \tag{102}$$

We define

$$\mathfrak{T}_\alpha = (-s_\alpha)^{-\frac{1}{2}}\mathfrak{H}_\alpha \qquad (\alpha = 2, \cdots, h); \tag{103}$$

then, by (67) and (102),

$$\mathfrak{T}'_\alpha = \bar{\mathfrak{T}}_\alpha > 0, \quad \mathfrak{T}_\alpha = \bar{\mathfrak{T}}_\alpha{}^{-1}\{\mathfrak{G}_\alpha\}, \quad \mathfrak{T}_\alpha = |s_\alpha|^{-1}\mathfrak{T}_\alpha{}^{-1}\{\mathfrak{H}'_\alpha\} \qquad (\alpha = 1, \cdots, h), \tag{104}$$

where $|s_\alpha|$ denotes the absolute value of s_α. The matrices $\mathfrak{T}_2, \cdots, \mathfrak{T}_h$ are fixed, whereas $\mathfrak{T}_1$ depends upon the variable point $\mathfrak{Z}$ of H, by (77). The space H is mapped onto a surface T, of $n(n+1)$ dimensions, in the space R. If $\mathfrak{W}$ is any element of the group Λ, then the transformation $\mathfrak{T}_\alpha \to \mathfrak{T}_\alpha\{\mathfrak{W}_\alpha{}^{-1}\}$ $(\alpha = 1, \cdots, h)$ maps T onto itself, and this mapping is the identical one, if and only if $\mathfrak{W} = \pm\mathfrak{E}$; on the other hand, by Lemma 6, the corresponding mapping in H is a symplectic transformation of $\mathfrak{Z}$, with the matrix $\mathfrak{M} = \mathfrak{C}^{-1}\mathfrak{W}\mathfrak{C}$ of the group Δ.

For any point $\mathfrak{T}$ of T, there exists a unimodular matrix $\mathfrak{U} = \mathfrak{U}_\mathfrak{Z}$ in K_0, such that $\mathfrak{T}\{\mathfrak{U}\}$ is a point of the domain R. By Lemma 10, we may choose a matrix $\mathfrak{L}$ of the finite set L, such that $\hat{\mathfrak{T}} = \mathfrak{T}\{\mathfrak{U}\mathfrak{L}\}$ belongs to $Q(\tau_4)$. Putting $\hat{\mathfrak{G}} = \mathfrak{G}[\mathfrak{U}\mathfrak{L}]$, $\hat{\mathfrak{H}} = \mathfrak{H}\{\mathfrak{U}\mathfrak{L}\}$ we obtain, by (104),

(105) $$\hat{\mathfrak{T}} = \bar{\mathfrak{T}}^{-1}\{\hat{\mathfrak{G}}\}, \qquad \hat{\mathfrak{T}} = |\,s\,|^{-1}\hat{\mathfrak{T}}^{-1}\{\hat{\mathfrak{H}}'\};$$

we omit here and in the following formulae the index α which runs always from 1 to h.

Let $\mathfrak{V}$ denote the matrix of the linear transformation $x_k \to x_{2n-k+1}$ $(k = 1, \cdots, 2n)$. If $\hat{\mathfrak{T}} = \mathfrak{P}\{\mathfrak{Q}\}$ is the decomposition defined in Section **34**, then $\bar{\mathfrak{T}}^{-1}\{\mathfrak{V}\} = \mathfrak{P}^{-1}\{\mathfrak{V}\}\{\mathfrak{V}\bar{\mathfrak{Q}}'^{-1}\mathfrak{V}\}$, and consequently the points $|\,s\,|^{-1}\hat{\mathfrak{T}}^{-1}\{\mathfrak{V}\}$ and $\hat{\mathfrak{T}}^{-1}\{\mathfrak{V}\}$ belong to a domain $Q(\tau_5)$, where τ_5 depends only upon K_0, n and s. Moreover

$$|\,\mathfrak{V}\hat{\mathfrak{H}}'\,| = (-1)^n\,|\,\mathfrak{H}\,|\,\mathrm{abs}\,(\mathfrak{U}\mathfrak{Q})^2, \qquad |\,\mathfrak{V}\hat{\mathfrak{G}}\,| = (-1)^n\,|\,\mathfrak{G}\,|\;|\,\mathfrak{U}\mathfrak{Q}\,|^2;$$

hence the norms of $|\,\mathfrak{V}\hat{\mathfrak{H}}'\,|$ and $|\,\mathfrak{V}\hat{\mathfrak{G}}\,|$ belong to a finite set. It follows now from Lemma 11 and (105), that also the matrices $\hat{\mathfrak{G}}$ and $\hat{\mathfrak{H}}$ belong to a finite set, independent of $\mathfrak{Z}$, and the same holds good for $\mathfrak{G}[\mathfrak{U}] = \hat{\mathfrak{G}}[\mathfrak{L}^{-1}]$ and $\mathfrak{H}\{\mathfrak{U}\} = \hat{\mathfrak{H}}\{\mathfrak{L}^{-1}\}$.

Choose now a complete system of points $\mathfrak{Z}_0$ in H, such that the pairs $\mathfrak{G}[\mathfrak{U}_0]$, $\mathfrak{H}\{\mathfrak{U}_0\}$ with $\mathfrak{U}_0 = \mathfrak{U}_{\mathfrak{Z}_0}$ are all different, and let V be the finite set of the unimodular matrices $\mathfrak{U}_0$. We denote by $G(\mathfrak{U}_0)$ the closed set of all points $\mathfrak{Z}$ of H, such that the corresponding point $\mathfrak{T}$ of T lies in $R_{\mathfrak{U}_0^{-1}}$, and by G the union of these $G(\mathfrak{U}_0)$, as $\mathfrak{U}_0$ runs over the elements of the set V.

Let $\mathfrak{Z}$ be again an arbitrary point of H and $\mathfrak{U} = \mathfrak{U}_{\mathfrak{Z}}$. Then there exists a uniquely determined $\mathfrak{U}_0$ in V, such that $\mathfrak{G}[\mathfrak{U}_0] = \mathfrak{G}[\mathfrak{U}]$ and $\mathfrak{H}\{\mathfrak{U}_0\} = \mathfrak{H}\{\mathfrak{U}\}$; thus $\mathfrak{U}_0\mathfrak{U}^{-1} = \mathfrak{W}$ is an element of the group Λ. Since $\mathfrak{T}\{\mathfrak{U}\}$ lies in R, the point $\mathfrak{T}\{\mathfrak{W}^{-1}\} = \mathfrak{T}\{\mathfrak{U}\mathfrak{U}_0^{-1}\}$ is contained in $R_{\mathfrak{U}_0^{-1}}$; hence $\mathfrak{Z}$ is mapped by the element $\mathfrak{M} = \mathfrak{C}^{-1}\mathfrak{W}\mathfrak{C}$ of Δ into a point of $G(\mathfrak{U}_0)$. This proves that any $\mathfrak{Z}$ in H is under Δ equivalent to at least one point of G; we call this point a *reduced image* of $\mathfrak{Z}$.

36. Let us assume that there exists in H a compact domain B, such that $\mathfrak{T}\{\mathfrak{U}_{\mathfrak{Z}}\}$ is a boundary point of R, relative to P, for all $\mathfrak{Z}$ in B. By Section **34**, the unimodular matrix $\mathfrak{U}_{\mathfrak{Z}}$ belongs then to a finite set, and we infer from (77), (100), (101), (103), that the expression $\mathfrak{p}'\mathfrak{S}\bar{\mathfrak{q}} + \bar{\mathfrak{p}}'\mathfrak{S}\mathfrak{q}$ has a constant value in H, where $\mathfrak{p} = \mathfrak{C}^{-1}\mathfrak{a}$ and $\mathfrak{q} = \mathfrak{C}^{-1}\mathfrak{b}$ with two columns $\mathfrak{a} \neq 0$ and $\mathfrak{b} \neq 0$ in K_0; moreover $\mathfrak{a} \neq \lambda\mathfrak{b}$ for every pure imaginary scalar factor λ. Replacing $\mathfrak{Y}$ by $\mathfrak{E} + \mathfrak{Y}$, we have the Taylor series

$$\mathfrak{S} = \mathfrak{E} + \mathfrak{S}_1 + \mathfrak{S}_2 + \cdots, \qquad \mathfrak{S}_1 = \begin{pmatrix} -\mathfrak{Y} & -\mathfrak{X} \\ -\mathfrak{X} & \mathfrak{Y} \end{pmatrix}, \qquad \mathfrak{S}_2 = \begin{pmatrix} \mathfrak{Y}^2 & \mathfrak{Y}\mathfrak{X} \\ \mathfrak{X}\mathfrak{Y} & \mathfrak{X}^2 \end{pmatrix},$$

in the neighborhood of $\mathfrak{Y} = 0$. It follows that the real part of $\mathfrak{p}'\mathfrak{S}_2\bar{\mathfrak{q}}$ vanishes identically in the real symmetric matrices $\mathfrak{X}$ and $\mathfrak{Y}$. Since $\mathfrak{a} = \mathfrak{C}\mathfrak{p}$, $\mathfrak{b} = \mathfrak{C}\mathfrak{q}$

and $\mathfrak{a} \neq \lambda\mathfrak{b}$, for all pure imaginary values of λ, we find easily $\mathfrak{a} = \omega\mathfrak{C}\mathfrak{r}$, $\mathfrak{b} = i\omega\mathfrak{C}\mathfrak{s}$ with real $\mathfrak{r}, \mathfrak{s}$ and complex scalar $\omega \neq 0$.

On the other hand $\mathfrak{G}[\mathfrak{C}] = \mathfrak{J}$, $\mathfrak{H}\{\mathfrak{C}\} = \sqrt{-s}\mathfrak{J}$, whence $\bar{\mathfrak{C}} = \sqrt{-s}\mathfrak{H}^{-1}\mathfrak{G}\mathfrak{C}$. We obtain, therefore,

$$\bar{\omega}\sqrt{-s}\mathfrak{H}^{-1}\mathfrak{G}\mathfrak{a} = \omega\bar{\mathfrak{a}}.$$

But $\mathfrak{G}$, $\mathfrak{H}$ and $\mathfrak{a} \neq 0$ are in $K_0 = K(\sqrt{-r})$, hence

$$\bar{\omega}\sqrt{-s} = \omega(\xi + \eta\sqrt{-r}); \qquad s = \xi^2 + r\eta^2, \tag{106}$$

where ξ, η are numbers in K. Since $r_\alpha > 0$, $s_\alpha < 0$ $(\alpha = 2, \cdots, h)$, the relationship (106) is only possible for $h = 1$; then K is the field of rational numbers.

It will be proved in Chapter IX, that $\Delta(\mathfrak{G}, \mathfrak{H})$ is commensurable with the modular group Γ, if and only if the diophantine equation (106) has a rational solution ξ, η. In this case, however, the construction and the properties of a fundamental domain for Δ follow in a simple way from the results of Chapter VI. Therefore we exclude this case for the rest of the present chapter.

For every point $\mathfrak{Z}_0$ of H, there exists now a unimodular matrix $\mathfrak{U}$ and a sequence of points $\mathfrak{Z}$ tending to $\mathfrak{Z}_0$, such that the corresponding points $\mathfrak{T}$ of T are inner points of $R_{\mathfrak{U}^{-1}}$, relative to P.

37. We denote by $F(\mathfrak{U}_0)$ the closure of the set of inner points of $G(\mathfrak{U}_0)$, relative to H. If $\mathfrak{Z}_0$ is a point of $G(\mathfrak{U}_0)$ which does not belong to $F(\mathfrak{U}_0)$, then we use the result of the last section to construct a sequence of points $\mathfrak{Z}$ tending to $\mathfrak{Z}_0$, such that they have as reduced images inner points of $G(\mathfrak{U})$, where $\mathfrak{U}$ is a certain matrix of the set V. Consequently $\mathfrak{Z}_0$ has a reduced image in $F(\mathfrak{U})$, and any point of H is equivalent under Δ to at least one point in one of the domains $F(\mathfrak{U}_0)$.

Let $\mathfrak{Z}^*$ be the image of $\mathfrak{Z}$ under the transformation $\mathfrak{M} = \mathfrak{C}^{-1}\mathfrak{W}\mathfrak{C}$ of Δ and let $\mathfrak{T}^*, \mathfrak{T}$ be the corresponding points in T. We assume now the existence of an inner point $\mathfrak{Z}$ of $F(\mathfrak{U}_0)$, such that $\mathfrak{Z}^*$ is a point of $F(\mathfrak{U}^*_0)$, where $\mathfrak{U}^*_0$ is also one of the matrices of the set V. By Section **36**, this holds then even under the further condition that $\mathfrak{T}\{\mathfrak{U}_0\}$ be an inner point of R, relative to P. But $\mathfrak{T}^*\{\mathfrak{U}^*_0\} = \mathfrak{T}\{\mathfrak{U}_0\}\{\mathfrak{U}_0^{-1}\mathfrak{W}^{-1}\mathfrak{U}^*_0\}$ is again a point of R, and consequently $\mathfrak{U}^*_0 = \pm\mathfrak{W}\mathfrak{U}_0$, $\mathfrak{G}[\mathfrak{U}^*_0] = \mathfrak{G}[\mathfrak{U}_0]$, $\mathfrak{H}\{\mathfrak{U}^*_0\} = \mathfrak{H}\{\mathfrak{U}_0\}$. Now it follows from the definition of V, that $\mathfrak{U}^*_0 = \mathfrak{U}_0$, $\mathfrak{W} = \pm\mathfrak{E}$. This proves that $F(\mathfrak{U}_0)$ and $F(\mathfrak{U}^*_0)$ do not overlap, if $\mathfrak{U}_0 \neq \mathfrak{U}^*_0$, and that the sum of the domains $F(\mathfrak{U}_0)$ is a fundamental domain F of Δ.

Obviously every $F(\mathfrak{U}_0)$ is bounded by a finite number of algebraic sur-

faces, and F has the same property. It follows immediately from Section **34**, that any compact domain in H is covered by a finite number of images $F_{\mathfrak{M}}$ of F under Δ and that F has only a finite number of neighbors $F_{\mathfrak{M}}$.

38. Let $\mathfrak{Z}$ be a point of $F(\mathfrak{U}_0)$ and $\mathfrak{S}$ the corresponding point of S. By (101) and (105),

$$\mathfrak{P} = s^{-1}\mathfrak{P}^{-1}\{\hat{\mathfrak{H}}'\{\bar{\mathfrak{Q}}^{-1}\}\}, \tag{107}$$

where

$$\begin{cases} \hat{\mathfrak{H}} = \mathfrak{H}\{\mathfrak{U}_0\mathfrak{Q}\}, \qquad \mathfrak{P}\{\mathfrak{Q}\} = \mathfrak{T}_1\{\mathfrak{U}_0\mathfrak{Q}\} = \mathfrak{S}\{\mathfrak{C}^{-1}\mathfrak{U}_0\mathfrak{Q}\} \\ \mathfrak{P} = [p_1, \cdots, p_{2n}], \quad 0 < p_k \leq \tau_4 p_{k+1} \qquad (k = 1, \cdots, 2n-1) \\ \mathfrak{Q} = (q_{kl}), \quad q_{kl} = 0 \ (k > l), \quad q_{kk} = 1, \quad \text{abs}\, q_{kl} \leq \tau_4 \quad (k < l). \end{cases} \tag{108}$$

Let d be the first diagonal element of $\hat{\mathfrak{H}}$. Then

$$p_1^2 \geq s^{-1}d^2 \tag{109}$$

and

$$\prod_{k=1}^{2n} p_k = (-s)^{-n} \mid \hat{\mathfrak{H}} \mid, \tag{110}$$

by (107).

We assume now $h > 1$, i. e., K is not the field of rational numbers. Then the conjugates of $\hat{\mathfrak{H}}$ in $K^{(\alpha)}(\sqrt{-r_\alpha})$ $(\alpha = 2, \cdots, h)$ are positive, and consequently $d \neq 0$. On the other hand, $\mathfrak{U}_0$ and $\mathfrak{Q}$ belong to a finite set. It follows from (108), (109), (110), that $\mathfrak{S}$ is bounded for all $\mathfrak{Z}$ in the fundamental domain F. By (77), the matrices $\mathfrak{Y}^{-1}$, $\mathfrak{Y} + \mathfrak{Y}^{-1}[\mathfrak{X}]$, $\mathfrak{Y}^{-1}\mathfrak{X}$ are bounded, hence also $\mathfrak{Y}$, $\mid \mathfrak{Y} \mid^{-1}$, $\mathfrak{X}$. This proves that F is compact.

In the remaining case $h = 1$, F is not necessarily compact, and the proof of the convergence of the volume integral $V(\Delta)$ requires more detailed estimates. This proof may be given by the same method which leads to the analogous result in the theory of units of indefinite quadratic forms; we omit it here.

It is also not difficult to prove that the space H is compact relative to $\Delta(\mathfrak{G}, \mathfrak{H})$, in the case $h = 1$, if and only if $n = 1$ and Δ is not commensurable with the elliptic modular group, and then also F is compact.

39. The congruence subgroup $\Lambda_\kappa(\mathfrak{G}, \mathfrak{H})$ of $\Lambda(\mathfrak{G}, \mathfrak{H})$ consists of all elements of Λ satisfying $\mathfrak{W} \equiv \mathfrak{E} \pmod{\kappa}$, where κ is a given ideal in K_0. Let ρ be a prime ideal of K_0 and p the rational prime number which is divisible by ρ; let κ be a power of ρ, such that p is not divisible by κ^{p-1}. If the transformation with the matrix $\mathfrak{C}^{-1}\mathfrak{W}\mathfrak{C} \neq \pm \mathfrak{E}$ has a fixed point in H, then $\mathfrak{W}$ is of finite order, since Δ is discontinuous in H.

We shall prove that $\Lambda_\kappa(\mathfrak{G}, \mathfrak{H})$ contains no element of finite order except $\mathfrak{E}$. Otherwise we may assume

$$\mathfrak{W}^q = \mathfrak{E}, \quad \mathfrak{W} = \mathfrak{E} + \mathfrak{R}, \quad \mathfrak{R} \equiv 0 \pmod{\rho^a}, \quad \mathfrak{R} \not\equiv 0 \pmod{\rho^{a+1}},$$

where q is a rational prime number and ρ^a divisible by κ. Then

$$\sum_{l=1}^{q} \binom{q}{l} \mathfrak{R}^l = 0; \quad \mathfrak{R} \equiv -q^{-1}\mathfrak{R}^q \pmod{\rho^{2a}}.$$

Since

$$q^{-1}\mathfrak{R}^q \equiv 0 \pmod{\rho^{a+1}},$$

we arrive at a contradiction.

The proofs of Theorems 9 and 10 are finished.

VIII. THE VOLUME OF THE FUNDAMENTAL DOMAIN OF THE MODULAR GROUP.

40. In the interval $0 \leq x \leq 1$, we consider an arbitrary monotone function $f(x)$, such that

$$f(1) = 0, \qquad \int_0^1 f(x) x^{n-1} dx = 1; \tag{111}$$

an example is $f(x) = n(n+1)(1-x)$. For $x \geq 1$, we put $f(x) = 0$. Let $\mathfrak{Z}$ be a point of H and $\mathfrak{S}$ the positive symmetric matrix defined in (77). For any $\epsilon > 0$, we define

$$\phi(\epsilon, \mathfrak{Z}) = \epsilon^{2n} \sum_{\mathfrak{w} \neq 0} f(\mathfrak{S}[\epsilon \mathfrak{w}]), \tag{112}$$

where $\mathfrak{w}$ runs over all lattice points $\neq 0$; this is a finite sum.

LEMMA 12. *If ϵ tends to 0 through positive values and $\mathfrak{Z}$ is fixed, then*

$$\lim \phi(\epsilon, \mathfrak{Z}) = \frac{\pi^n}{(n-1)!}.$$

On account of the definition of the integral, we have

$$\lim_{\epsilon \to 0} \phi(\epsilon, \mathfrak{Z}) = \int_{\mathfrak{S}[\mathfrak{q}] \leq 1} f(\mathfrak{S}[\mathfrak{q}]) d\mathfrak{q},$$

where $d\mathfrak{q}$ denotes the euclidean volume element in the space of the real vectors $\mathfrak{q}$ of $2n$ dimensions. Since the volume of the ellipsoid $\mathfrak{S}[\mathfrak{q}] \leq x$ is

$$J(x) = \frac{\pi^n}{n!} x^n \qquad (x \geq 0),$$

we obtain

$$\int_{\mathfrak{S}[\mathfrak{q}]\leq 1} f(\mathfrak{S}[\mathfrak{q}])\,d\mathfrak{q} = \int_0^1 f(x)\,\frac{dJ(x)}{dx}\,dx = \frac{\pi^n}{(n-1)!}\int_0^1 f(x)x^{n-1}dx = \frac{\pi^n}{(n-1)!}\cdot$$

41. Let F be the fundamental domain of the modular group, in the space H of $\mathfrak{Z} = \mathfrak{X} + i\mathfrak{Y}$, and denote by dv the euclidean volume element in the space of $\mathfrak{X}$ and $\mathfrak{Y}^{-1}$, of $n(n+1)$ dimensions.

LEMMA 13. *There exists an integrable positive function $g(\mathfrak{Z})$ of the elements of $\mathfrak{Z}$, independent of ϵ, such that*

$$\phi(\epsilon, \mathfrak{Z}) \leq g(\mathfrak{Z}) \qquad (0 < \epsilon \leq 1) \tag{113}$$

and the integral

$$\gamma = \int_F g(\mathfrak{Z})\,dv$$

converges.

We denote by $h(\rho, \mathfrak{Z})$ the number of lattice points $\mathfrak{w}$ satisfying $\mathfrak{S}[\mathfrak{w}] \leq \rho$, where ρ is an arbitrary positive number. By (78),

$$\mathfrak{S}[\mathfrak{w}] = \mathfrak{Y}^{-1}[\mathfrak{u} - \mathfrak{X}\mathfrak{v}] + \mathfrak{Y}[\mathfrak{v}], \qquad \mathfrak{w} = \begin{pmatrix}\mathfrak{u}\\ \mathfrak{v}\end{pmatrix},$$

and consequently $h(\rho, \mathfrak{Z})$ is not larger than the number of integral solutions $\mathfrak{u}, \mathfrak{v}$ of

$$\mathfrak{Y}^{-1}[\mathfrak{u} - \mathfrak{X}\mathfrak{v}] \leq \rho, \qquad \mathfrak{Y}[\mathfrak{v}] \leq \rho. \tag{114}$$

Put $\mathfrak{Y}^{-1} = \mathfrak{P}[\mathfrak{Q}]$ with $\mathfrak{P} = [p_1, \cdots, p_n]$ and $\mathfrak{Q} = (q_{kl})$, $q_{kl} = 0$ $(k > l)$, $q_{kk} = 1$, and let u_k, v_k, r_k $(k = 1, \cdots, n)$ denote the elements of the columns $\mathfrak{u}, \mathfrak{v}, \mathfrak{r} = \mathfrak{X}\mathfrak{v}$. The first condition (114) involves

$$p_k\{(u_k - r_k) + \sum_{l=k+1}^{n} q_{kl}(u_l - r_l)\}^2 \leq \rho \qquad (k = 1, \cdots, n);$$

this proves that the number of integral $\mathfrak{u}$ is

$$\leq \prod_{k=1}^{n} (1 + 2p_k^{-\frac{1}{2}}\rho^{\frac{1}{2}}),$$

for any given $\mathfrak{v}$. On the other hand, the second condition (114) involves

$$p_k^{-1}(v_k + \sum_{l=1}^{k-1} q^*_{kl}v_l)^2 \leq \rho,$$

where $(q^*_{kl}) = \mathfrak{Q}^{-1}$; this proves that the number of integral $\mathfrak{v}$ is

$$\leq \prod_{k=1}^{n} (1 + 2p_k^{\frac{1}{2}}\rho^{\frac{1}{2}}).$$

It follows that

$$h(\rho, \mathfrak{Z}) \leq \prod_{k=1}^{n} (1 + 2\rho^{\frac{1}{2}}(p_k^{\frac{1}{2}} + p_k^{-\frac{1}{2}}) + 4\rho).$$

Let $\gamma_1, \cdots, \gamma_4$ denote positive numbers which depend only upon n. By (93) and (94), we have $p_k < \gamma_1$ $(k = 1, \cdots, n)$, for all $\mathfrak{Z}$ in F; hence

$$h(\rho, \mathfrak{Z}) < \gamma_2(1 + \rho)^n \prod_{k=1}^{n} p_k^{-\frac{1}{2}}. \tag{115}$$

Now $0 \leq f(x) \leq f(0)$ $(0 \leq x \leq 1)$ and $f(x) = 0$ $(x \geq 1)$; consequently we infer from the definition (112) that

$$\phi(\epsilon, \mathfrak{Z}) \leq \epsilon^{2n} f(0) h(\epsilon^{-2}, \mathfrak{Z}). \tag{116}$$

By (115) and (116),

$$\phi(\epsilon, \mathfrak{Z}) < f(0)\gamma_2(1 + \epsilon^2)^n \prod_{k=1}^{n} p_k^{-\frac{1}{2}} < f(0)\gamma_3 \prod_{k=1}^{n} p_k^{-\frac{1}{2}} \qquad (0 < \epsilon \leq 1).$$

We define

$$g(\mathfrak{Z}) = f(0)\gamma_3 \prod_{k=1}^{n} p_k^{-\frac{1}{2}};$$

then (113) is fulfilled and it remains only to prove the convergence of the integral

$$\gamma_4 = \int_F \prod_{k=1}^{n} p_k^{-\frac{1}{2}} dv.$$

Instead of the elements Y_{kl} $(k \leq l)$ of $(Y_{kl}) = \mathfrak{Y}^{-1} = \mathfrak{P}[\mathfrak{Q}]$, we introduce the new variables $p_k^{\frac{1}{2}}$ $(k = 1, \cdots, n)$ and q_{kl} $(k < l)$. The functional determinant has the value

$$2^n \prod_{k=1}^{n} p_k^{n-k+\frac{1}{2}},$$

whence

$$\gamma_4 = 2^n \int_F \prod_{k=1}^{n} p_k^{n-k} dv_1$$

with

$$dv_1 = \prod_{k \leq l} dx_{kl} \prod_{k<l} dq_{kl} \prod_{k} dp_k^{\frac{1}{2}}.$$

Since $\mathfrak{X}$, $\mathfrak{P}$, $\mathfrak{Q}$ are bounded in F, the convergence is obvious.

Applying a well-known theorem of integral calculus, we obtain from Lemmata 12 and 13 the important

LEMMA 14. *The integral*

$$\psi(\epsilon) = \int_F \phi(\epsilon, \mathfrak{Z})\, dv = \epsilon^{2n} \sum_{\mathfrak{w} \neq 0} \int_F f(\mathfrak{S}[\mathfrak{w}])\, dv$$

converges and

$$\lim_{\epsilon \to 0} \psi(\epsilon) = \frac{\pi^n}{(n-1)!} \int_F dv = \frac{\pi^n}{(n-1)!} V_n$$

where $V_n = V(\Gamma_n)$ *is the symplectic volume of the fundamental domain for the modular group* $\Gamma = \Gamma_n$ *of degree* n.

42. We denote by W the set of all columns $\mathfrak{w}^{(2n)}$ with coprime integral elements and by $\mathfrak{e}$ the first column of the unit matrix $\mathfrak{E}^{(2n)}$; obviously $\mathfrak{e}$ belongs to W.

LEMMA 15. *There exists a modular matrix* $\mathfrak{M}$ *with the first column* $\mathfrak{w}$, *if and only if* $\mathfrak{w}$ *belongs to* W.

The necessity of the condition is obvious, since $\mathfrak{M}$ has integral elements and the determinant 1. Therefore we have only to prove the existence of a modular matrix $\mathfrak{M}$ satisfying $\mathfrak{M}^{-1}\mathfrak{w} = \mathfrak{e}$, where $\mathfrak{w}$ is a given column of the set W.

Consider the column

$$\mathfrak{M}^{-1}\mathfrak{w} = \mathfrak{w}_1 = \begin{pmatrix} \mathfrak{u}^{(n)} \\ \mathfrak{v}^{(n)} \end{pmatrix},$$

where $\mathfrak{M}$ is an arbitrary modular matrix, and let u, v denote the greatest common divisors of the n elements of $\mathfrak{u}, \mathfrak{v}$; we define $u = 0$ or $v = 0$ if $\mathfrak{u} = 0$ or $\mathfrak{v} = 0$. We choose now the matrix $\mathfrak{M}$, such that the sum $u + v = w$ is as small as possible. If $u < v$, we replace $\mathfrak{M}, \mathfrak{u}, \mathfrak{v}$ by $\mathfrak{M}\mathfrak{J}, -\mathfrak{v}, \mathfrak{u}$ and obtain the case $u > v$, with the same value of w; hence we may assume $u \geq v$.

If $v > 0$, we determine a unimodular matrix $\mathfrak{U}$ with the first row $v^{-1}\mathfrak{v}'$ and an integral column $\mathfrak{t}$, such that all elements of the column $\mathfrak{U}\mathfrak{u} - v\mathfrak{t} = \mathfrak{u}_1$ have an absolute value $\leq (v/2)$; then the greatest common divisor u_1 of these elements satisfies $u_1 \leq (v/2) < u$. Let $\mathfrak{T}$ be an integral symmetric matrix having $\mathfrak{t}$ as its first column and

$$\mathfrak{M}_1 = \begin{pmatrix} \mathfrak{U}^{-1} & \mathfrak{U}^{-1}\mathfrak{T} \\ 0 & \mathfrak{U}' \end{pmatrix}.$$

Since $\mathfrak{U}'^{-1}\mathfrak{v} = v\mathfrak{e}_1$ and $\mathfrak{T}\mathfrak{U}'^{-1}\mathfrak{v} = v\mathfrak{t}$, where $\mathfrak{e}_1$ denotes the first column of the unit matrix $\mathfrak{E}^{(n)}$, we find

$$(\mathfrak{M}\mathfrak{M}_1)^{-1}\mathfrak{w} = \mathfrak{M}_1^{-1}\mathfrak{w}_1 = \begin{pmatrix} \mathfrak{u}_1 \\ v\mathfrak{e}_1 \end{pmatrix}.$$

But $\mathfrak{M}_1$ and $\mathfrak{M}\mathfrak{M}_1$ are again modular matrices and $u_1 + v < w$, in contradiction to the minimum property of w.

Consequently $v = 0$ and $u = 1$, the elements of $\mathfrak{w}_1$ being coprime. Let $\mathfrak{U}_1$ be a unimodular matrix with the first row $\mathfrak{u}'$ and

$$\mathfrak{M}_2 = \begin{pmatrix} \mathfrak{U}'_1 & 0 \\ 0 & \mathfrak{U}_1^{-1} \end{pmatrix}.$$

Then $(\mathfrak{M}\mathfrak{M}_2)^{-1}\mathfrak{w} = \mathfrak{e}$ and $\mathfrak{M}\mathfrak{M}_2$ is unimodular, q. e. d.

We denote by Δ_1 the subgroup of Γ consisting of all modular matrices $\mathfrak{M}$ with the first column $\mathfrak{e}$. Obviously two arbitrary modular matrices are then and only then in the same left coset of Δ_1 relative to the homogeneous modular group Γ_0, if they have the same first column. Applying Lemma 15, we find the decomposition

$$\Gamma_0 = \sum_{\mathfrak{w} \subset W} \mathfrak{M}_{\mathfrak{w}} \Delta_1 = \sum_{\mathfrak{w} \subset W} \Delta_1 \mathfrak{M}_{\mathfrak{w}}^{-1},$$

where $\mathfrak{w}$ runs over all elements of the set W and $\mathfrak{M}_{\mathfrak{w}}$ denotes a modular matrix with the first column $\mathfrak{w}$. We choose $\mathfrak{M}_{-\mathfrak{w}} = -\mathfrak{M}_{\mathfrak{w}}$, such that $\mathfrak{M}_{\mathfrak{w}}$ and $\mathfrak{M}_{-\mathfrak{w}}$ give the same element of Γ. Let $F_{\mathfrak{w}}$ be the image of F under the transformation $\mathfrak{M}_{\mathfrak{w}}^{-1}$; then the domains $F_{\mathfrak{w}}$ cover exactly twice a fundamental domain F_0 of the group Δ_1. On the other hand

$$\int_F f(\mathfrak{S}[\epsilon\mathfrak{w}])\,dv = \int_{F_{\mathfrak{w}}} f(\mathfrak{S}[\epsilon\mathfrak{e}]\,dv,$$

by Lemma 6. Using the abbreviation $p = \mathfrak{S}[\mathfrak{e}]$ for the first diagonal element of $\mathfrak{S}$, we obtain

$$\sum_{\mathfrak{w} \subset W} \int_F f(\mathfrak{S}[\epsilon\mathfrak{w}])\,dv = 2 \int_{F_0} f(\epsilon^2 p)\,dv.$$

Now we replace ϵ by ϵl and sum over all positive integers l; then $l\mathfrak{w}$ runs exactly over all lattice points $\neq 0$, and we have

$$\psi(\epsilon) = 2\epsilon^{2n} \sum_{l=1}^{\infty} \int_{F_0} f(\epsilon^2 l^2 p)\,dv, \tag{117}$$

where $\psi(\epsilon)$ is the function defined in Lemma 14.

43. Any element of Δ_1 has the form

$$\mathfrak{M} = \begin{pmatrix} \mathfrak{A} & \mathfrak{B} \\ \mathfrak{C} & \mathfrak{D} \end{pmatrix}, \quad \mathfrak{A} = \begin{pmatrix} 1 & * \\ 0 & \mathfrak{A}_1 \end{pmatrix}, \quad \mathfrak{B} = \begin{pmatrix} * & * \\ * & \mathfrak{B}_1 \end{pmatrix}, \tag{118}$$

$$\mathfrak{C} = \begin{pmatrix} 0 & * \\ 0 & \mathfrak{C}_1 \end{pmatrix}, \quad \mathfrak{D} = \begin{pmatrix} * & * \\ * & \mathfrak{D}_1 \end{pmatrix}.$$

It follows from (7) that also

$$\mathfrak{A}_1\mathfrak{B}'_1 = \mathfrak{B}_1\mathfrak{A}'_1, \quad \mathfrak{C}_1\mathfrak{D}'_1 = \mathfrak{D}_1\mathfrak{C}'_1, \quad \mathfrak{A}_1\mathfrak{D}'_1 - \mathfrak{B}_1\mathfrak{C}'_1 = \mathfrak{E};$$

hence

$$\mathfrak{M}_1 = \begin{pmatrix} \mathfrak{A}_1 & \mathfrak{B}_1 \\ \mathfrak{C}_1 & \mathfrak{D}_1 \end{pmatrix}$$

belongs to the homogeneous modular group of degree $n-1$. We define

$$\mathfrak{A}_2 = \begin{pmatrix} 1 & 0 \\ 0 & \mathfrak{A}_1 \end{pmatrix}, \quad \mathfrak{B}_2 = \begin{pmatrix} 0 & 0 \\ 0 & \mathfrak{B}_1 \end{pmatrix}, \quad \mathfrak{C}_2 = \begin{pmatrix} 0 & 0 \\ 0 & \mathfrak{C}_1 \end{pmatrix},$$

$$\mathfrak{D}_2 = \begin{pmatrix} 1 & 0 \\ 0 & \mathfrak{D}_1 \end{pmatrix}, \quad \mathfrak{M}_2 = \begin{pmatrix} \mathfrak{A}_2 & \mathfrak{B}_2 \\ \mathfrak{C}_2 & \mathfrak{D}_2 \end{pmatrix},$$

$$\mathfrak{M}_0 = \mathfrak{M}\mathfrak{M}_2^{-1} = \begin{pmatrix} \mathfrak{A} & \mathfrak{B} \\ \mathfrak{C} & \mathfrak{D} \end{pmatrix}\begin{pmatrix} \mathfrak{D}'_2 & -\mathfrak{B}'_2 \\ -\mathfrak{C}'_2 & \mathfrak{A}'_2 \end{pmatrix} = \begin{pmatrix} \mathfrak{A}_0 & \mathfrak{B}_0 \\ \mathfrak{C}_0 & \mathfrak{D}_0 \end{pmatrix},$$

where

$$\mathfrak{A}_0 = \begin{pmatrix} 1 & * \\ 0 & \mathfrak{E} \end{pmatrix}, \quad \mathfrak{B}_0 = \begin{pmatrix} * & * \\ * & 0 \end{pmatrix}, \quad \mathfrak{C}_0 = \begin{pmatrix} 0 & * \\ 0 & 0 \end{pmatrix}, \quad \mathfrak{D}_0 = \begin{pmatrix} * & * \\ * & \mathfrak{E} \end{pmatrix}.$$

Since $\mathfrak{M}_0$ is symplectic, we have $\mathfrak{C}_0\mathfrak{D}'_0 = \mathfrak{D}_0\mathfrak{C}'_0$, whence $\mathfrak{C}_0 = 0$; moreover $\mathfrak{A}_0\mathfrak{D}'_0 = \mathfrak{E}$ and $\mathfrak{A}_0\mathfrak{B}'_0 = \mathfrak{B}_0\mathfrak{A}'_0$ so that $\mathfrak{A}_0 = \mathfrak{U}'$, $\mathfrak{D}_0 = \mathfrak{U}^{-1}$, $\mathfrak{B}_0 = \mathfrak{T}\mathfrak{U}^{-1}$ with

$$\mathfrak{U} = \begin{pmatrix} 1 & 0 \\ \mathfrak{a} & \mathfrak{E} \end{pmatrix}, \qquad \mathfrak{T} = \begin{pmatrix} b & \mathfrak{b}' \\ \mathfrak{b} & 0 \end{pmatrix} \tag{119}$$

and integral $\mathfrak{a}, \mathfrak{b}, b$.

On the other hand, if $\mathfrak{U}$ and $\mathfrak{T}$ are defined by (119) with arbitrary integral $\mathfrak{a}, \mathfrak{b}, b$, then

$$\mathfrak{M}_0 = \begin{pmatrix} \mathfrak{U}' & \mathfrak{T}\mathfrak{U}^{-1} \\ 0 & \mathfrak{U}^{-1} \end{pmatrix} \tag{120}$$

is a modular matrix and $\mathfrak{M} = \mathfrak{M}_0\mathfrak{M}_2$ has again the form (118). It is obvious that the matrices $\mathfrak{M}_2$ constitute a group Δ_2 which is isomorphic to the homogeneous modular group of degree $n-1$ and is a subgroup of Δ_1. The left cosets of Δ_2 relative to Δ_1 are of the form $\mathfrak{M}_0\Delta_2$, where $\mathfrak{M}_0$ runs over all matrices defined by (119) and (120).

This result enables us to construct another fundamental domain of Δ_1. Let $\hat{\mathfrak{Z}} = (\mathfrak{A}_2\mathfrak{Z} + \mathfrak{B}_2)(\mathfrak{C}_2\mathfrak{Z} + \mathfrak{D}_2)^{-1}$ be the modular transformation with the matrix $\mathfrak{M}_2$. We decompose

$$\mathfrak{Z} = \begin{pmatrix} * & * \\ * & \mathfrak{Z}_1 \end{pmatrix}, \qquad \hat{\mathfrak{Z}} = \begin{pmatrix} * & * \\ * & \hat{\mathfrak{Z}}_1 \end{pmatrix}$$

and obtain

$$\mathfrak{A}_2\mathfrak{Z} + \mathfrak{B}_2 = \begin{pmatrix} * & * \\ * & \mathfrak{A}_1\mathfrak{Z}_1 + \mathfrak{B}_1 \end{pmatrix}, \qquad \mathfrak{C}_2\mathfrak{Z} + \mathfrak{D}_2 = \begin{pmatrix} 1 & 0 \\ * & \mathfrak{C}_1\mathfrak{Z}_1 + \mathfrak{D}_1 \end{pmatrix},$$

$$(\mathfrak{C}_2\mathfrak{Z} + \mathfrak{D}_2)^{-1} = \begin{pmatrix} 1 & 0 \\ * & (\mathfrak{C}_1\mathfrak{Z}_1 + \mathfrak{D}_1)^{-1} \end{pmatrix};$$

consequently $\hat{\mathfrak{Z}}_1 = (\mathfrak{A}_1\mathfrak{Z}_1 + \mathfrak{B}_1)(\mathfrak{C}_1\mathfrak{Z}_1 + \mathfrak{D}_1)^{-1}$ is the image of $\mathfrak{Z}_1$ under the modular transformation with the matrix $\mathfrak{M}_1$, of degree $n-1$. For any $\mathfrak{Z}$ in H, we determine $\mathfrak{M}_1$, such that $\hat{\mathfrak{Z}}_1$ lies in the fundamental domain F_1 of the modular group of degree $n-1$. Since we may replace $\mathfrak{M}_1$ by $-\mathfrak{M}_1$, there are always two different possibilities for the choice of the corresponding element $\mathfrak{M}_2$ of Δ_2.

We replace $\hat{\mathfrak{Z}}$ again by $\mathfrak{Z}$ and consider now the modular transformation $\hat{\mathfrak{Z}} = \mathfrak{Z}[\mathfrak{U}] + \mathfrak{T}$ with the matrix $\mathfrak{M}_0$. By (119),

$$\hat{\mathfrak{Z}}_1 = \mathfrak{Z}_1, \quad \hat{\mathfrak{Y}}^{-1} = \mathfrak{Y}^{-1}\begin{bmatrix} 1 & -\mathfrak{a}' \\ 0 & \mathfrak{E} \end{bmatrix}, \quad \hat{\mathfrak{x}} = \mathfrak{x}\begin{bmatrix} 1 & 0 \\ \mathfrak{a} & \mathfrak{E} \end{bmatrix} + \begin{pmatrix} b & \mathfrak{b}' \\ \mathfrak{b} & 0 \end{pmatrix},$$

where $\hat{\mathfrak{x}} + i\hat{\mathfrak{y}} = \hat{\mathfrak{Z}}$. On account of the definition (77), the matrices $\mathfrak{Y}^{-1}$ and $\mathfrak{S}$ have the same first diagonal element p. Putting $\mathfrak{Y}^{-1} = (Y_{kl})$ and $\hat{\mathfrak{Y}}^{-1} = (\hat{Y}_{kl})$, we obtain $\hat{Y}_{1l} = Y_{1l} - pa_l$ $(l = 2, \cdots, n)$, where $a_2, \cdots, a_n$ denote the elements of $\mathfrak{a}$. We determine first $\mathfrak{a}$, such that $-(p/2) \leq \hat{Y}_{1l} \leq p/2$ $(l = 2, \cdots, n)$, and then b, $\mathfrak{b}$, such that the elements $\hat{x}_{1l}$ of the first row of $\hat{\mathfrak{x}}$ satisfy $-\frac{1}{2} \leq \hat{x}_{1l} \leq \frac{1}{2}$ $(l = 1, \cdots, n)$.

It follows that any point of H has, relative to Δ_1, an equivalent point in the domain defined by the conditions

$$\mathfrak{Z}_1 \subset F_1, \qquad \begin{matrix} -(p/2) \leq Y_{1l} \leq p/2 & (l = 2, \cdots, n), \\ -\frac{1}{2} \leq x_{1l} \leq \frac{1}{2} & (l = 1, \cdots, n). \end{matrix} \tag{121}$$

On account of the ambiguity in the choice of $\mathfrak{M}_2$, this domain is not yet a fundamental domain of Δ_1. It is transformed into itself by the particular mapping $z_{1l} \to -z_{1l}$ $(l = 2, \cdots, n)$, $z_{11} \to z_{11}$, $\mathfrak{Z}_1 \to \mathfrak{Z}_1$, obtained from $\mathfrak{M}_1 = -\mathfrak{E}$, $\mathfrak{M}_0 = \mathfrak{E}$. By the additional condition

$$Y_{12} \geq 0 \tag{122}$$

together with (121), we obtain now a fundamental domain F^* of Δ_1.

In the special case $n = 1$, the condition (122) does not exist, and (121) reduces to $-\frac{1}{2} \leq x_{11} \leq \frac{1}{2}$.

44. By Lemma 6, the first element p of $\mathfrak{S}$ is invariant under all transformations of the group Δ_1. Since F_0 and F^* are both fundamental domains of Δ_1, we obtain

$$\int_{F_0} f(\epsilon^2 l^2 p)\, dv = \int_{F^*} f(\epsilon^2 l^2 p)\, dv. \tag{123}$$

We use now the decomposition

$$\mathfrak{Y}^{-1} = \begin{pmatrix} p^{-1} & 0 \\ 0 & \mathfrak{Y}_1^{-1} \end{pmatrix}\begin{bmatrix} p & \mathfrak{y}' \\ 0 & \mathfrak{E} \end{bmatrix} = \begin{pmatrix} p & \mathfrak{y}' \\ \mathfrak{y} & \mathfrak{Y}_0 \end{pmatrix},$$

where

$$\mathfrak{Y}_1^{-1} = \mathfrak{Y}_0 - p^{-1}\mathfrak{y}\mathfrak{y}' \tag{124}$$

and $\mathfrak{Y}_1$ denotes the imaginary part of $\mathfrak{Z}_1$. Introducing as new variables the elements η_{kl} $(1 \leq k \leq l \leq n-1)$ of $\mathfrak{Y}_1^{-1} = (\eta_{kl})$ instead of the elements Y_{kl} $(2 \leq k \leq l \leq n)$ of $\mathfrak{Y}_0$, we obtain, by (124),

$$\begin{aligned} dv &= \prod_{1 \leq k \leq l \leq n} (dx_{kl} dY_{kl}) \\ &= \prod_{2 \leq k \leq l \leq n} (dx_{kl} d\eta_{kl}) \prod_{l=1}^{n} (dx_{1l} dY_{1l}) = dv_1 \prod_{l=1}^{n} (dx_{1l} dY_{1l}), \end{aligned} \tag{125}$$

where dv_1 is the symplectic volume element for $\mathfrak{Z}_1$ instead of $\mathfrak{Z}$; moreover $Y_{11} = p$.

Define $\eta_n = 1$ for $n = 1$ and $= 2$ for $n > 1$. By (111), (121), (122) and (125),

$$\begin{aligned} \eta_n \int_{F^*} f(\epsilon^2 l^2 p)\, dv &= V_{n-1} \int_0^\infty p^{n-1} f(\epsilon^2 l^2 p)\, dp \\ &= V_{n-1} (\epsilon l)^{-2n} \int_0^1 x^{n-1} f(x)\, dx = V_{n-1} (\epsilon l)^{-2n}, \end{aligned} \tag{126}$$

where V_{n-1} is the volume $V(\Gamma_{n-1})$ of the fundamental domain F_1 for the modular group of degree $n-1$ and $V_0 = 1$. By (117), (123) and (126),

$$\psi(\epsilon) = 2\eta_n^{-1} V_{n-1} \zeta(2n),$$

independent of ϵ. Lemma 14 leads now to the recursion formula

$$V_n = 2\eta_n^{-1} (n-1)!\, \pi^{-n} \zeta(2n)\, V_{n-1},$$

whence

$$V_n = 2 \prod_{k=1}^{n} \{(k-1)!\, \pi^{-1} \zeta(2k)\}, \tag{127}$$

and Theorem 11 is proved.

45. By a well-known result of Euler,

$$2(2\pi)^{-2k} \zeta(2k) = \frac{B_{2k}}{(2k)!} \qquad (k = 1, 2, \cdots),$$

where B_{2k} is the absolute value of the Bernoulli numbers

$$B_2 = \frac{1}{6}, \quad -B_4 = -\frac{1}{30}, \quad B_6 = \frac{1}{42}, \quad -B_8 = -\frac{1}{30}, \cdots.$$

Thus we obtain the expression

$$V_n = 2^{n^2+1} \pi^{n(n+1)/2} \prod_{k=1}^{n} \left\{ \frac{(k-1)!}{(2k)!}\, B_{2k} \right\}$$

and in particular

$$V_1 = \frac{\pi}{3}, \quad V_2 = \frac{\pi^3}{270}, \quad V_3 = \frac{\pi^6}{127575}, \quad V_4 = \frac{\pi^{10}}{200930625}.$$

Consider now the congruence subgroup $\Gamma_n(p)$ of the modular group of degree n, defined by $\mathfrak{M} \equiv \mathfrak{E} \pmod{p}$, where p is a prime number. Its index has the value

$$j_n(p) = \prod_{k=1}^{n} p^{2k-1}(p^{2k}-1). \tag{128}$$

For $p \neq 2$, this group has no fixed point in H, by Theorem 10. *If* Theorem 5 still holds good for the non-compact fundamental domain of $\Gamma_n(p)$, then this open manifold has the Euler number

$$\chi = c_n(-\pi)^{-n(n+1)/2} V_n j_n(p).$$

We denote by $\chi_n(p)$ the right-hand side of this hypothetic formula. Using the values of c_1, c_2, c_3 given in Theorem 5, we find

$$c_1 V_1 = -\frac{\pi}{3!}, \qquad c_2 V_2 = -\frac{\pi^3}{6!}, \qquad c_3 V_3 = \frac{\pi^6}{2 \cdot 9!}$$

and consequently

$$\chi_1(p) = -\frac{1}{6} p(p^2-1), \quad \chi_2(p) = -\frac{1}{720} p^4(p^2-1)(p^4-1),$$
$$\chi_3(p) = \frac{1}{2 \cdot 9!} p^9(p^2-1)(p^4-1)(p^6-1).$$

It is easily proved that these rational numbers are integers for all prime numbers p; in particular $\chi_1(2) = -1$, $\chi_2(2) = -1$, $\chi_3(2) = 2$.

46. Let $q = p^a$ $(a \geq 1)$ be a power of a prime number p. For any integral skew-symmetric matrix $\mathfrak{G}^{(2n)}$, we denote by $A_q(\mathfrak{G})$ the number of integral solutions $\mathfrak{L}$, modulo q, of the congruence

$$\mathfrak{F}[\mathfrak{L}] \equiv \mathfrak{G} \pmod{q}. \tag{129}$$

In particular, $A_q(\mathfrak{F}) = E_q$ is the order of the homogeneous modular group modulo q, namely the number of incongruent solutions $\mathfrak{M}$ of

$$\mathfrak{F}[\mathfrak{M}] \equiv \mathfrak{F} \pmod{q}. \tag{130}$$

It is known that $E_p = j_n(p)$ has the value (128), and more generally

$$E_q = q^{n(2n+1)} \prod_{k=1}^{n} (1 - p^{-2k}).$$

By (129) and (130), $\mathfrak{M}\mathfrak{L}$ is also a solution of (129); we call it equiva-

lent to $\mathfrak{L}$, relative to the group of the $\mathfrak{M}$. Since there are exactly E_q matrices in each class of equivalent solutions, the number of different classes of solutions of (129) is $A_q(\mathfrak{G})/E_q$. On the other hand, the number of incongruent skew-symmetric matrices $\mathfrak{G}$ is $q^{n(2n-1)}$, and

$$\sum_{\mathfrak{G}} A_q(\mathfrak{G}) = q^{4n^2}$$

is the number of incongruent $\mathfrak{L}$. The average number of classes of solutions is therefore

$$q^{-n(2n-1)} \sum_{\mathfrak{G}} \frac{A_q(\mathfrak{G})}{E_q} = \prod_{k=1}^{n} (1 - p^{-2k})^{-1} = d_p, \tag{131}$$

independent of the exponent a.

By Euler's formula

$$\zeta(s) = \prod_p (1 - p^{-s})^{-1}; \qquad (s > 1),$$

the result (127) may be written

$$d_0 = \prod_p d_p,$$

where p runs over all prime numbers and

$$d_0 = \tfrac{1}{2} V_n \prod_{k=1}^{n} \frac{\pi^k}{(k-1)!}. \tag{132}$$

Now the main formula in the analytic theory of quadratic forms suggests that d_0 can be defined independently as a density connected with the real solutions $\mathfrak{L}$ of the equation $\mathfrak{F}[\mathfrak{L}] = \mathfrak{G}$. We shall prove that d_0 has the value defined in (4); this is the statement of Theorem 12.

47. Let

$$\mathfrak{G} = \begin{pmatrix} \mathfrak{G}_1 & \mathfrak{G}_2 \\ -\mathfrak{G}'_2 & \mathfrak{G}_3 \end{pmatrix}$$

be a real skew-symmetric matrix and

$$\mathfrak{F} = \begin{pmatrix} \mathfrak{E} & \mathfrak{E} \\ -i\mathfrak{E} & i\mathfrak{E} \end{pmatrix}, \qquad \frac{1}{2i}\mathfrak{G}\{\mathfrak{F}\} = \begin{pmatrix} \mathfrak{H}_1 & -\bar{\mathfrak{A}} \\ \mathfrak{A} & -\bar{\mathfrak{H}}_1 \end{pmatrix} = \mathfrak{K}_1; \tag{133}$$

then $\mathfrak{H}_1 = \frac{1}{2}(\mathfrak{G}_2 + \mathfrak{G}'_2) + \frac{1}{2i}(\mathfrak{G}_1 + \mathfrak{G}_3)$ is hermitian and $\mathfrak{A} = \frac{1}{2}(\mathfrak{G}_2 - \mathfrak{G}'_2) + \frac{1}{2i}(\mathfrak{G}_1 - \mathfrak{G}_3)$ is complex skew-symmetric. For $\mathfrak{G} = \mathfrak{J}$, we have $\mathfrak{H}_1 = \mathfrak{E}$ and $\mathfrak{A} = 0$. We choose a neighborhood G of $\mathfrak{G} = \mathfrak{J}$, such that $|\mathfrak{H}_1| \neq 0$ and that the characteristic roots $b_1, \cdots, b_n$ of the matrix

$$\mathfrak{B} = -\bar{\mathfrak{A}}\bar{\mathfrak{H}}_1^{-1}\mathfrak{A}\mathfrak{H}_1^{-1} \tag{134}$$

are of absolute value < 1, for all $\mathfrak{G} \subset G$.

The series

$$\mathfrak{W} = \sum_{k=0}^{\infty} \binom{\frac{1}{2}}{k} \mathfrak{B}^k \tag{135}$$

converges and satisfies the equation

$$\mathfrak{W}^2 = \mathfrak{E} + \mathfrak{B}.$$

The characteristic roots of the matrix $\mathfrak{E} + \mathfrak{W}$ are $1 + (1 + b_k)^{\frac{1}{2}} \neq 0$ $(k = 1, \cdots, n)$. We define

$$\mathfrak{H} = \tfrac{1}{2}(\mathfrak{E} + \mathfrak{W})\mathfrak{H}_1; \tag{136}$$

then $|\mathfrak{H}| \neq 0$ and $|2\mathfrak{H} - \mathfrak{H}_1| \neq 0$.

LEMMA 16. *The matrix $\mathfrak{H}$ is hermitian and*

$$\mathfrak{H} + \tfrac{1}{4}\bar{\mathfrak{A}}\bar{\mathfrak{H}}^{-1}\mathfrak{A} = \mathfrak{H}_1. \tag{137}$$

We have

$$\mathfrak{B}\bar{\mathfrak{A}} = -\bar{\mathfrak{A}}\bar{\mathfrak{H}}_1^{-1}\mathfrak{A}\mathfrak{H}_1^{-1}\bar{\mathfrak{A}} = \bar{\mathfrak{A}}\mathfrak{B}' \quad \text{and} \quad \mathfrak{B}\mathfrak{H}_1 = -\bar{\mathfrak{A}}\bar{\mathfrak{H}}_1^{-1}\mathfrak{A} = \mathfrak{H}_1\mathfrak{B}',$$

whence

$$\bar{\mathfrak{H}}' = \tfrac{1}{2}\mathfrak{H}_1(\mathfrak{E} + \bar{\mathfrak{W}}') = \tfrac{1}{2}(\mathfrak{E} + \mathfrak{W})\mathfrak{H}_1 = \mathfrak{H}$$

and

$$\begin{aligned}\tfrac{1}{2}\bar{\mathfrak{A}}\bar{\mathfrak{H}}^{-1}\mathfrak{A} = \bar{\mathfrak{A}}(\mathfrak{E} + \mathfrak{W}')\bar{\mathfrak{H}}_1^{-1}\mathfrak{A} &= -(\mathfrak{E} + \mathfrak{W})^{-1}\mathfrak{B}\mathfrak{H}_1 \\ &= (\mathfrak{E} - \mathfrak{W})\mathfrak{H}_1 = 2\mathfrak{H}_1 - 2\mathfrak{H},\end{aligned}$$

q. e. d.

We define now

$$\mathfrak{K} = \begin{pmatrix} \mathfrak{H} & 0 \\ 0 & -\bar{\mathfrak{H}} \end{pmatrix}, \qquad \mathfrak{C} = \begin{pmatrix} \mathfrak{E} & -\frac{1}{2}\bar{\mathfrak{H}}^{-1}\mathfrak{A} \\ -\frac{1}{2}\mathfrak{H}^{-1}\bar{\mathfrak{A}} & \mathfrak{E} \end{pmatrix}$$

and obtain, by (133) and (137), $\mathfrak{K}\{\mathfrak{C}\} = \mathfrak{K}_1$ and

$$\mathfrak{C}\begin{pmatrix} \mathfrak{E} & \frac{1}{2}\bar{\mathfrak{H}}^{-1}\mathfrak{A} \\ \frac{1}{2}\mathfrak{H}^{-1}\bar{\mathfrak{A}} & \mathfrak{E} \end{pmatrix} = \begin{pmatrix} 2\mathfrak{E} - \bar{\mathfrak{H}}^{-1}\bar{\mathfrak{H}}_1 & 0 \\ 0 & 2\mathfrak{E} - \mathfrak{H}^{-1}\mathfrak{H}_1 \end{pmatrix},$$

whence in particular $|\mathfrak{C}| \neq 0$.

48. We consider now the set L of real matrices $\mathfrak{L}$ satisfying $\mathfrak{J}[\mathfrak{L}] = \mathfrak{G} \subset G$. Putting

$$\mathfrak{L}\mathfrak{F} = \mathfrak{V}_1 = \begin{pmatrix} \mathfrak{P}_1 & \bar{\mathfrak{P}}_1 \\ \mathfrak{Q}_1 & \bar{\mathfrak{Q}}_1 \end{pmatrix}, \qquad \mathfrak{V}_1\mathfrak{C}^{-1} = \begin{pmatrix} \mathfrak{P} & \bar{\mathfrak{P}} \\ \mathfrak{Q} & \bar{\mathfrak{Q}} \end{pmatrix} = \mathfrak{V}, \tag{138}$$

we have

$$\frac{1}{2i}\mathfrak{J}\{\mathfrak{V}\} = \frac{1}{2i}\mathfrak{G}\{\mathfrak{F}\mathfrak{C}^{-1}\} = \mathfrak{K}_1\{\mathfrak{C}^{-1}\} = \mathfrak{K}$$

and consequently

$$\frac{1}{2i}(\mathfrak{P}'\bar{\mathfrak{Q}} - \mathfrak{Q}'\bar{\mathfrak{P}}) = \mathfrak{H}, \qquad \mathfrak{P}'\mathfrak{Q} - \mathfrak{Q}'\mathfrak{P} = 0; \tag{139}$$

moreover

$$\mathfrak{P}_1 = \mathfrak{P} - \tfrac{1}{2}\bar{\mathfrak{P}}\mathfrak{H}^{-1}\bar{\mathfrak{A}}, \qquad \mathfrak{Q}_1 = \mathfrak{Q} - \tfrac{1}{2}\bar{\mathfrak{Q}}\mathfrak{H}^{-1}\bar{\mathfrak{A}}.$$

By (133), (134), (135), (136) and (137), the neighborhood G of $\mathfrak{G} = \mathfrak{J}$ is mapped onto a neighborhood G^* of $\mathfrak{H} = \mathfrak{E}$, $\mathfrak{A} = 0$ in the $(\mathfrak{H}, \mathfrak{A})$ space. Since $|\mathfrak{H}| \neq 0$, we have $\mathfrak{H} > 0$. By (139), $|\mathfrak{Q}| \neq 0$ and

$$\mathfrak{P}\mathfrak{Q}^{-1} = \mathfrak{Z} = \mathfrak{X} + i\mathfrak{Y}$$

is a point of H with the imaginary part $\mathfrak{Y} = \mathfrak{H}\{\mathfrak{Q}^{-1}\}$. Then

$$(140) \qquad \mathfrak{P}_1 = (\mathfrak{Z} - \tfrac{1}{2}\bar{\mathfrak{Z}}\mathfrak{Y}^{-1}\bar{\mathfrak{A}}[\mathfrak{Q}^{-1}])\mathfrak{Q}, \quad \mathfrak{Q}_1 = (\mathfrak{E} - \tfrac{1}{2}\mathfrak{Y}^{-1}\bar{\mathfrak{A}}[\mathfrak{Q}^{-1}])\mathfrak{Q}.$$

It follows, from (138) and (140), that L is mapped onto the set L^* of the $(\mathfrak{Z}, \mathfrak{Q}, \mathfrak{A})$ space defined by the conditions $\mathfrak{Z} \subset H$, $(\mathfrak{Y}\{\mathfrak{Q}\}, \mathfrak{A}) \subset G^*$.

If

$$\mathfrak{M} = \begin{pmatrix} \mathfrak{M}_1 & \mathfrak{M}_2 \\ \mathfrak{M}_3 & \mathfrak{M}_4 \end{pmatrix}$$

is any modular matrix, then the mapping

$$(141) \qquad \mathfrak{Z} \to (\mathfrak{M}_1\mathfrak{Z} + \mathfrak{M}_2)(\mathfrak{M}_3\mathfrak{Z} + \mathfrak{M}_4)^{-1}, \qquad \mathfrak{Q} \to (\mathfrak{M}_3\mathfrak{Z} + \mathfrak{M}_4)\mathfrak{Q}$$

transforms L^* into itself and leaves every point of G^* invariant. We restrict now $\mathfrak{Z}$ to the fundamental domain F of the modular group Γ. Since the particular mapping with $\mathfrak{M} = -\mathfrak{E}$ leaves $\mathfrak{Z}$ invariant and changes $\mathfrak{Q}$ into $-\mathfrak{Q}$, we obtain in L^* a fundamental domain L^*_0 for all mappings (141), if we impose on $\mathfrak{Q}$ a linear homogeneous condition, e. g., $\sigma(\mathfrak{Q} + \bar{\mathfrak{Q}}) \geq 0$; then L^*_0 is defined by $\mathfrak{Z} \subset F$, $(\mathfrak{Y}\{\mathfrak{Q}\}, \mathfrak{A}) \subset G^*$, $\sigma(\mathfrak{Q} + \bar{\mathfrak{Q}}) \geq 0$. Let L_0 be the corresponding domain in L. Obviously L_0 is a fundamental domain in L relative to the homogeneous modular group Γ_0, such that the images of L_0 under the mappings $\mathfrak{L} \to \mathfrak{M}\mathfrak{L}$ cover L completely without gaps and overlappings.

Denote by $v(L_0)$ the euclidean volume of L_0, the elements of $\mathfrak{L}$ being considered as rectangular cartesian coördinates, and by $v(G)$ the euclidean volume of G, where the elements of the skew-symmetric matrix $\mathfrak{G} = \mathfrak{J}[\mathfrak{L}]$ above the diagonal are the coördinates. Theorem 12 asserts that

$$(142) \qquad \lim_{G \to \mathfrak{J}} \frac{v(L_0)}{v(G)} = d_0,$$

if G runs over any sequence of neighborhoods tending to the single point $\mathfrak{J}$, with the value d_0 defined in (132). Obviously the left-hand side of (142) is the analogue of the expression in (131), for the real valuation instead of the p-adic valuation.

49. In order to simplify the notation, we introduce cartesian coördinates for a matrix $\mathfrak{T}^{(m)} = (t_{kl})$ in the following way: If $\mathfrak{T}$ is arbitrarily real, we choose the m^2 coördinates t_{kl} $(k, l = 1, \cdots, m)$; if $\mathfrak{T}$ is symmetric, we take the $\frac{1}{2}m(m+1)$ coördinates t_{kl} $(1 \leq k \leq l \leq m)$; if $\mathfrak{T}$ is skew-symmetric, we take the $\frac{1}{2}m(m-1)$ coördinates t_{kl} $(1 \leq k < l \leq m)$. If $\mathfrak{T}$ is complex, we split

$$\mathfrak{T} = \mathfrak{T}_1 + i\mathfrak{T}_2, \qquad \mathfrak{T}_1 = \tfrac{1}{2}(\mathfrak{T} + \bar{\mathfrak{T}}), \qquad \mathfrak{T}_2 = \frac{1}{2i}(\mathfrak{T} - \bar{\mathfrak{T}}),$$

and proceed in the same manner with the real part $\mathfrak{T}_1$ and the imaginary part $\mathfrak{T}_2$. In particular, a hermitian matrix $\mathfrak{T}$ has then the m^2 coördinates

$$t_{kk}\ (k = 1, \cdots, m), \quad \tfrac{1}{2}(t_{kl} + \bar{t}_{kl}), \quad \frac{1}{2i}(t_{kl} - \bar{t}_{kl}) \qquad (1 \leq k < l \leq m).$$

In all these cases, we denote by $d\mathfrak{T}$ the euclidean volume element in the space of the coördinates of $\mathfrak{T}$.

By (133), (138) and (140), we have

$$v(L_0) = \int_{L_0} d\mathfrak{Q}, \qquad d\mathfrak{Q} = d\mathfrak{P}_1 d\mathfrak{Q}_1 = |\,\mathfrak{Q}_1\bar{\mathfrak{Q}}_1\,|^n\, d\mathfrak{R}\, d\mathfrak{Q}_1$$

with

$$\mathfrak{R} = \mathfrak{P}_1\mathfrak{Q}_1^{-1} = (\mathfrak{Z} + \tfrac{1}{2}\bar{\mathfrak{Z}}\mathfrak{Y}^{-1}\bar{\mathfrak{A}}[\mathfrak{Q}^{-1}])\,(\mathfrak{E} - \tfrac{1}{2}\mathfrak{Y}^{-1}\bar{\mathfrak{A}}[\mathfrak{Q}^{-1}])^{-1}.$$

In a sufficiently small neighborhood of $\mathfrak{A} = 0$, the power series

$$\mathfrak{R} = \mathfrak{Z} - i\bar{\mathfrak{A}}[\mathfrak{Q}^{-1}] + \cdots$$

converges, whence

$$d\mathfrak{R} \sim 2^{n(n-1)}\,|\,\mathfrak{Q}\bar{\mathfrak{Q}}\,|^{1-n}\, d\mathfrak{Z}\, d\mathfrak{A}, \qquad d\mathfrak{Q}_1 \sim d\mathfrak{Q} \qquad (\mathfrak{A} \to 0).$$

It follows that

$$v(L_0) \sim 2^{n(n-1)} \int_{L^*_0} |\,\mathfrak{Q}\bar{\mathfrak{Q}}\,|\, d\mathfrak{X}\, d\mathfrak{Y}\, d\mathfrak{Q}\, d\mathfrak{A} \qquad (G \to \mathfrak{J}).$$

We choose a real matrix $\mathfrak{C}_1^{(n)}$, such that $\mathfrak{C}'_1\mathfrak{C}_1 = \mathfrak{Y}$. Then the condition $\mathfrak{Y}\{\mathfrak{Q}\} = \mathfrak{H}$ is replaced by $\mathfrak{Q}'_2\bar{\mathfrak{Q}}_2 = \mathfrak{H}$ with $\mathfrak{Q}_2 = \mathfrak{C}_1\mathfrak{Q}$. Since

$$d\mathfrak{Q}_2 = |\,\mathfrak{C}_1\,|^{2n}\, d\mathfrak{Q} = |\,\mathfrak{Y}\,|^n\, d\mathfrak{Q}, \qquad |\,\mathfrak{Q}\bar{\mathfrak{Q}}\,| = |\,\mathfrak{Q}_2\bar{\mathfrak{Q}}_2\,|\;|\,\mathfrak{Y}\,|^{-1},$$

we obtain

$$(143)\quad v(L_0) \sim 2^{n(n-1)} \int_F |\,\mathfrak{Y}\,|^{-n-1}\, d\mathfrak{X}\, d\mathfrak{Y} \cdot \tfrac{1}{2} \int_{(\mathfrak{Q}'_2\bar{\mathfrak{Q}}_2,\,\mathfrak{A}) \subset G^*} |\,\mathfrak{Q}_2\bar{\mathfrak{Q}}_2\,|\, d\mathfrak{Q}_2\, d\mathfrak{A} \quad (G \to \mathfrak{J}).$$

On the other hand, by (133) and (136),

$$v(G) = \int_G d\mathfrak{G}, \qquad d\mathfrak{G} = d\mathfrak{G}_1 d\mathfrak{G}_2 d\mathfrak{G}_3 = 2^{n(n-1)} d\mathfrak{H}_1 d\mathfrak{A},$$
$$d\mathfrak{H}_1 \sim d\mathfrak{H} \qquad (G \to \mathfrak{Z});$$

hence

$$v(G) \sim 2^{n(n-1)} \int_{(\mathfrak{H}, \mathfrak{A}) \subset G^*} d\mathfrak{H} d\mathfrak{A}. \tag{144}$$

The first integral in (143) has the value $V_n = V(\Gamma)$. By (132), (143) and (144), the proof of (142) is reduced to the proof of the following lemma.

LEMMA 17. *Let H_1 be a domain in the space of the positive hermitian matrices and $\mathfrak{H}^* > 0$; then*

$$\int_{\mathfrak{H}^*\{\mathfrak{Q}\} \subset H_1} d\mathfrak{Q} = c_n \mid \mathfrak{H}^* \mid^{-n} \int_{\mathfrak{H} \subset H_1} d\mathfrak{H}, \qquad c_n = \prod_{k=1}^{n} \frac{\pi^k}{(k-1)!}. \tag{145}$$

We determine a matrix $\mathfrak{C}^*$ satisfying $\mathfrak{H}^*\{\mathfrak{C}^*\} = \mathfrak{E}$ and replace $\mathfrak{Q}$, $d\mathfrak{Q}$ by $\mathfrak{C}^*\mathfrak{Q}$, $\mid \mathfrak{C}^*\bar{\mathfrak{C}}^* \mid^n d\mathfrak{Q}$. Then we have only to prove (145) in the special case $\mathfrak{H}^* = \mathfrak{E}$.

We apply induction and assume first $n > 1$. Let $\mathfrak{Q}_0$ be a matrix having the same first column $\mathfrak{q}$ as $\mathfrak{Q}$ and $\mid \mathfrak{Q}_0 \mid \neq 0$; then

$$\mathfrak{Q} = \mathfrak{Q}_0 \begin{pmatrix} 1 & \mathfrak{t}' \\ 0 & \mathfrak{T} \end{pmatrix},$$
$$d\mathfrak{Q} = \mid \mathfrak{Q}_0 \bar{\mathfrak{Q}}_0 \mid^{n-1} d\mathfrak{q} d\mathfrak{t} d\mathfrak{T}.$$

Introducing

$$\mathfrak{Q}'_0 \bar{\mathfrak{Q}}_0 = \mathfrak{H}_0 = \begin{pmatrix} h & \bar{\mathfrak{d}}'_0 \\ \mathfrak{d}_0 & \mathfrak{D}_0 \end{pmatrix} = \begin{pmatrix} h & 0 \\ 0 & \mathfrak{N}_0 \end{pmatrix} \begin{Bmatrix} 1 & h^{-1}\mathfrak{d}'_0 \\ 0 & \mathfrak{E} \end{Bmatrix},$$

$$\mathfrak{Q}' \cdot \bar{\mathfrak{Q}} = \mathfrak{H} = \begin{pmatrix} h & \bar{\mathfrak{d}}' \\ \mathfrak{d} & \mathfrak{D} \end{pmatrix} = \begin{pmatrix} h & 0 \\ 0 & \mathfrak{N} \end{pmatrix} \begin{Bmatrix} 1 & h^{-1}\mathfrak{d}' \\ 0 & \mathfrak{E} \end{Bmatrix}, \tag{146}$$

we have

$$h = \mathfrak{q}'\bar{\mathfrak{q}}, \quad \mathfrak{N} = \mathfrak{D} - h^{-1}\mathfrak{d}\bar{\mathfrak{d}}',$$
$$h \mid \mathfrak{N}_0 \mid = \mid \mathfrak{H}_0 \mid = \mid \mathfrak{Q}_0 \bar{\mathfrak{Q}}_0 \mid, \quad \mathfrak{N} = \mathfrak{N}_0\{\mathfrak{T}\}, \quad \mathfrak{d} = \mathfrak{T}'\mathfrak{d}_0 + h\mathfrak{t}. \tag{147}$$

We replace $\mathfrak{t}$ by the new variable $\mathfrak{d}$ and obtain

$$d\mathfrak{t} = h^{2-2n} d\mathfrak{d},$$
$$d\mathfrak{Q} = \mid \mathfrak{Q}_0 \bar{\mathfrak{Q}}_0 \mid^{n-1} h^{2-2n} d\mathfrak{q} d\mathfrak{d} d\mathfrak{T}.$$

Using (145) with $n-1$, $\mathfrak{N}_0$, $\mathfrak{T}$, $\mathfrak{N}$ instead of n, $\mathfrak{H}^*$, $\mathfrak{Q}$, $\mathfrak{H}$, we find

$$\int_{\mathfrak{Q}'\bar{\mathfrak{Q}} \subset H_1} d\mathfrak{Q} = c_{n-1} \int \mid \mathfrak{Q}_0 \bar{\mathfrak{Q}}_0 \mid^{n-1} h^{2-2n} \mid \mathfrak{N}_0 \mid^{1-n} d\mathfrak{q} d\mathfrak{d} d\mathfrak{N}, \tag{148}$$

where the domain of integration for the variables $\mathfrak{q}$, $\mathfrak{d}$, $\mathfrak{N}$ is given by (146), (147) and the condition $\mathfrak{H} \subset H_1$. We take the new variable $\mathfrak{D}$ instead of $\mathfrak{N}$, with $d\mathfrak{D} = d\mathfrak{N}$, and perform the integration over $\mathfrak{q}$, for fixed values of $\mathfrak{d}$ and $\mathfrak{D}$. Now

$$\int_{\mathfrak{q}'\bar{\mathfrak{q}} \leq g} d\mathfrak{q} = J(g) = \frac{\pi^n}{n!} g^n; \qquad (g > 0),$$

whence

$$\int_{h_1 \leq h \leq h_2} h^{1-n} d\mathfrak{q} = \int_{h_1}^{h_2} h^{1-n} \frac{dJ(h)}{dh}\, dh = \frac{\pi^n}{(n-1)!} \int_{h_1}^{h_2} dh. \tag{149}$$

But

$$|\, \mathfrak{Q}_0 \bar{\mathfrak{Q}}_0 \,|^{n-1} h^{2-2n} \,|\, \mathfrak{N}_0 \,|^{1-n} = h^{1-n}, \qquad c_{n-1} \frac{\pi^n}{(n-1)!} = c_n$$

and consequently, by (148) and (149),

$$\int_{\mathfrak{Q}'\bar{\mathfrak{Q}} \subset H_1} d\mathfrak{Q} = c_n \int_{\mathfrak{H} \subset H_1} d\mathfrak{H}. \tag{150}$$

It remains to prove (150) in the case $n = 1$. Then it is contained in formula (149) which holds good also in this case.

50. The relationship

$$d_0 = \prod_p d_p$$

constitutes another example of the "integral formulae" of analytic number theory which appear in the theory of class fields and in the theory of quadratic forms. The formula also holds good in the case of an arbitrary group $\Delta(\mathfrak{G}, \mathfrak{H})$, if the densities d_0 and d_p are defined in an analogous manner. But we do not go into detail since the proof of this statement depends upon the analytic theory of hermitian forms, of which no complete account has hitherto been given.

We proved in Section **31**, that the matrix $\mathfrak{Y}^{-1}$ is bounded in the fundamental domain F of the modular group. Theorem 11 may be used for an estimation of the maximum μ_n of $|\, \mathfrak{Y} \,|^{-1}$ in F. It follows from (80), (90) and (91), that F is contained in the domain F_1 defined by the conditions

$$|\, \mathfrak{Y} \,|^{-1} \leq \mu_n, \quad \mathfrak{Y}^{-1} \subset R, \quad -\tfrac{1}{2} \leq x_{kl} \leq \tfrac{1}{2} \qquad (1 \leq k \leq l \leq n),$$

where R is the Minkowski domain of Section **29**. Consequently the symplectic volume of F_1 is equal to the euclidean volume of that part of R which is defined by $|\, \mathfrak{X} \,| \leq \mu_n$, $\mathfrak{X} \subset R$. By an important result of Minkowski, this volume is

$$\frac{2}{n+1} \mu_n^{(n+1)/2} \prod_{k=2}^{n} \{\pi^{-k/2} \Gamma\left(\frac{k}{2}\right) \zeta(k)\}. \tag{151}$$

Hence this number is an upper bound of V_n. Using (127), we obtain the estimation

$$(152) \qquad \mu_n \geq \nu_n, \qquad \nu_n^{(n+1)/2} = \frac{n+1}{6}\,\pi \prod_{k=2}^{n} \left\{ \pi^{-k/2}\,\frac{\Gamma(k)\zeta(2k)}{\Gamma(k/2)\zeta(k)} \right\}$$

and in particular

$$\mu_1 \geq \frac{\pi}{3}, \qquad \mu_2^{3/2} \geq \frac{\pi^2}{30}, \qquad \mu_3^{\,2} \geq \frac{8\pi^6}{42525\zeta(3)},$$

the exact value of μ_1 being $2/3^{1/2}$. By Stirling's formula, for $n \to \infty$,

$$\log \nu_n = \frac{n}{2}\left(\log \frac{2n}{\pi} - \frac{3}{2}\right) + O(1); \quad \log \mu_n > \frac{n}{2}\log n + O(n).$$

This proves that μ_n increases rapidly, as a function of n.

Let now $\mathfrak{Z} = \mathfrak{X} + i\mathfrak{Y}$ be an arbitrary point of H and let $\mathfrak{Z}^* = \mathfrak{X}^* + i\mathfrak{Y}^* = (\mathfrak{A}\mathfrak{Z} + \mathfrak{B})(\mathfrak{C}\mathfrak{Z} + \mathfrak{D})^{-1}$ be the image of $\mathfrak{Z}$ under any modular transformation. By Section **30**, the expression

$$|\,\mathfrak{Y}^*\,|^{-1} = |\,\mathfrak{Y}\,|^{-1}\,\mathrm{abs}\,(\mathfrak{C}\mathfrak{Z} + \mathfrak{D})^2$$

has its minimum value, if and only if $\mathfrak{Z}^*$ lies in F. Hence μ_n is the maximum of these minima, for the set of all $\mathfrak{Z}$ in H. A fortiori, there exists a matrix $\mathfrak{Z}$ in H, such that the inequality

$$\mathrm{abs}\,(\mathfrak{C}\mathfrak{Z} + \mathfrak{D})^2 \geq \nu_n\,|\,\mathfrak{Y}\,|$$

holds for all second matrix rows $(\mathfrak{C}\mathfrak{D})$ of modular matrices, where ν_n is the number defined in (152). Writing $\mathfrak{Z} = \mathfrak{P}\mathfrak{Q}^{-1}$, we have

$$\mathfrak{P}'\mathfrak{Q} - \mathfrak{Q}'\mathfrak{P} = 0, \quad \frac{1}{2i}\,(\mathfrak{P}'\bar{\mathfrak{Q}} - \mathfrak{Q}'\bar{\mathfrak{P}}) > 0$$

and

$$\mathrm{abs}\,(\mathfrak{C}\mathfrak{P} + \mathfrak{D}\mathfrak{Q})^2 \geq \nu_n \left| \frac{1}{2i}\,(\mathfrak{P}'\bar{\mathfrak{Q}} - \mathfrak{Q}'\bar{\mathfrak{P}}) \right|.$$

IX. COMMENSURABLE GROUPS.

51. We proved in Chapter V that the group $\Delta_0(\mathfrak{G}, \mathfrak{H})$ is commensurable with $\Delta_0(r, s)$. In order to demonstrate Theorem 13, we have now to discuss the necessary and sufficient conditions for the commensurability of two groups $\Delta_0(r, s)$ and $\Delta_0(r_1, s_1)$. We assumed that the numbers r and s are integers of the totally real field K; all conjugates of s except s itself are negative, whereas r is totally positive. Obviously s generates the field K. The numbers r_1 and s_1 have the same properties with respect to the totally real field K_1.

With the groups $\Delta_0(r,s)$ and $\Delta_0(r_1,s_1)$ we associate the two quadratic forms

$$Q(r,s) = rsx^2 - ry^2 + sz^2 \quad \text{and} \quad Q(r_1,s_1) = r_1s_1x_1^2 - r_1y_1^2 + s_1z_1^2$$

in 3 variables.

LEMMA 18. *If $Q(r,s)$ and $Q(r_1,s_1)$ are equivalent in K, then $K = K_1$ and the groups $\Delta_0(r,s)$, $\Delta_0(r_1,s_1)$ are commensurable.*

Under the assumption of the lemma, s_1 is certainly a number of K; therefore $K_1 \subset K$. Since all conjugates of $Q(r,s)$ are negative definite except $Q(r,s)$ itself, the same holds good for $Q(r_1,s_1)$ considered as a quadratic form in K; consequently s_1 is also a generating number of K and $K = K_1$.

Using the quaternion units $\epsilon_0, \epsilon_1, \epsilon_2, \epsilon_3$ defined in (74), we have $Q(r,s)\epsilon_0 = (x\epsilon_1 + y\epsilon_2 + z\epsilon_3)^2$. Let $\mathfrak{F} = (f_{kl})$ be the matrix of the linear substitution transforming $Q(r,s)$ into $Q(r_1,s_1)$ and let

$$\eta_k = \sum_{l=1}^{3} f_{lk}\epsilon_l \qquad (k = 1,2,3).$$

Then

$$\eta_1^2 = r_1s_1\epsilon_0, \quad \eta_2^2 = -r_1\epsilon_0, \quad \eta_3^2 = s_1\epsilon_0, \quad \eta_k\eta_l = -\eta_l\eta_k \quad (1 \le k < l \le 3)$$

and consequently there exists a real matrix $\mathfrak{Q}^{(2)}$ satisfying

$$|\mathfrak{Q}| = 1, \qquad \mathfrak{Q}^{-1}\eta_k\mathfrak{Q} = \omega_k \qquad (k = 1,2,3),$$

$$\omega_1 = \sqrt{r_1s_1}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad \omega_2 = \pm\sqrt{r_1}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad \omega_3 = \pm\sqrt{s_1}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

The elements of $\Delta_0(r,s)$ are defined by

$$\mathfrak{M} = \sum_{k=0}^{3} \mathfrak{A}_k \times \epsilon_k, \qquad \mathfrak{F}[\mathfrak{M}] = \mathfrak{F},$$

with integral $\mathfrak{A}_k$ $(k = 0, \cdots, 3)$ in K. Putting

$$(153) \qquad (\mathfrak{E}^{(n)} \times \mathfrak{F}^{-1})(\mathfrak{A}_1\mathfrak{A}_2\mathfrak{A}_3)' = (\mathfrak{B}_1\mathfrak{B}_2\mathfrak{B}_3)', \qquad \mathfrak{E}^{(n)} \times \mathfrak{Q} = \mathfrak{K},$$

we obtain

$$\mathfrak{M} = \mathfrak{A}_0 \times \epsilon_0 + \sum_{k=1}^{3} \mathfrak{B}_k \times \eta_k = \mathfrak{K}\mathfrak{M}_1\mathfrak{K}^{-1}$$

with

$$(154) \qquad \mathfrak{M}_1 = \mathfrak{A}_0 \times \epsilon_0 + \sum_{k=1}^{3} \mathfrak{B}_k \times \omega_k,$$

and $\mathfrak{K}$ is symplectic.

Let q be an even positive rational integer, such that the matrix $q\mathfrak{F}^{-1}/2rs$ is integral. If $\mathfrak{M} \equiv \mathfrak{E} \pmod{q}$, then $\mathfrak{A}_0 + \sqrt{rs}\mathfrak{A}_1 \equiv \mathfrak{E}$, $\mathfrak{A}_0 - \sqrt{rs}\mathfrak{A}_1 \equiv \mathfrak{E}$,

$\sqrt{r}\mathfrak{A}_2 + \sqrt{s}\mathfrak{A}_3 \equiv 0$, $-\sqrt{r}\mathfrak{A}_2 + \sqrt{s}\mathfrak{A}_3 \equiv 0$; hence $\mathfrak{A}_0$ is integral and $\mathfrak{A}_k \equiv 0 \pmod{q/2rs}$ $(k = 1, 2, 3)$. By (153) and (154), $\mathfrak{B}_k$ is also integral and $\mathfrak{M}_1 = \mathfrak{R}^{-1}\mathfrak{M}\mathfrak{R}$ is an element of the group $\Delta_0(r_1, s_1)$. Consider now in $\Delta_0(r, s)$ the subgroup of all $\mathfrak{M}$, such that $\mathfrak{M}_1$ is contained in $\Delta_0(r_1, s_1)$. Since this subgroup Δ^*_0 contains the congruence subgroup of $\Delta_0(r, s)$, for the module q, it is of finite index in $\Delta_0(r, s)$. On the other hand, $\mathfrak{R}^{-1}\Delta^*_0\mathfrak{R}$ consists of all $\mathfrak{M}_1$ in $\Delta_0(r_1, s_1)$, such that $\mathfrak{M} = \mathfrak{R}\mathfrak{M}_1\mathfrak{R}^{-1}$ is an element of $\Delta_0(r, s)$; hence the same argument shows that $\mathfrak{R}^{-1}\Delta^*_0\mathfrak{R}$ is a subgroup of finite index in $\Delta_0(r_1, s_1)$. It follows that $\Delta_0(r, s)$ and $\Delta_0(r_1, s_1)$ are commensurable.

52. On account of Lemma 18, the condition of Theorem 13 is sufficient for the commensurability of $\Delta_0(r, s)$ and $\Delta_0(r_1, s_1)$. It remains to prove that this condition is also necessary, which is a little more difficult.

LEMMA 19. *Let a and b be two numbers of K, $ab \neq 0$ and K_0 an arbitrary algebraic number field. There exists an integral number t in K, such that $at^2 - b$ is not the square of a number of K_0 and $a(at^2 - b)$ is totally positive.*

We may obviously assume that a, b are integral and $K \subset K_0$. We choose in K a prime ideal λ having the following 3 properties: λ is not divisible by the square of a prime ideal of K_0, i. e., λ is not a factor of the relative discriminant of K_0 with respect to K; $2ab$ is not divisible by λ; ab is a quadratic residue modulo λ in K. Then there exists an integer t in K, such that $at^2 - b$ is divisible by λ, but not by λ^2, and that, moreover, $a(at^2 - b)$ is totally positive. Since $at^2 - b$ is divisible by exactly the first power of a prime ideal in K_0, it cannot be a square number in K_0.

LEMMA 20. *There exists a quadratic form $Q(r_0, s_0)$ equivalent with $Q(r, s)$ in K and a quadratic form $Q(r_2, s_2)$ equivalent with $Q(r_1, s_1)$ in K_1, such that the field $K^*(\sqrt{r_0}, \sqrt{s_0}, \sqrt{r_2}, \sqrt{s_2})$ has the degree 16 relative to K^*, where K^* is the union of the fields K and K_1.*

Using Lemma 19, we choose an integral number t in K, such that $rt^2 - s = r_0$ is totally positive and no square in K^*. Then the quadratic form $Q(r, s) = rsx^2 - ry^2 + sz^2$ is equivalent with $rsx^2 - r_0y^2 + r_0rsz^2$ in K. Applying again Lemma 19, we construct an integer u in K, such that $r_0rsu^2 + rs = s_0$ is no square in $K^*(\sqrt{r_0})$ and ss_0 is totally positive. Then $Q(r, s)$ is equivalent with $Q(r_0, s_0)$ in K and the field $K^*(\sqrt{r_0}, \sqrt{s_0})$ has the

degree 4 relative to K^*. We complete the proof of the lemma by using the same argument for $Q(r_1, s_1)$, K_1, $K^*(\sqrt{r_0}, \sqrt{s_0})$ instead of $Q(r, s)$, K, K^*.

By Lemmata 18 and 20, we can assume for the rest of the proof of Theorem 13, that $K^*(\sqrt{r}, \sqrt{s}, \sqrt{r_1}, \sqrt{s_1})$ has the degree 16 relative to the union K^* of K and K_1.

LEMMA 21. *Let G_1 be a subgroup of another group G, of finite index j. For any element A of G, there exists a positive rational integer $g \leq j$, such that A^g is an element of G_1.*

Let $G_1, G_2, \cdots, G_j$ be the right cosets of G_1 relative to G and consider the $j+1$ powers A^k $(k = 0, \cdots, j)$. Then two of these powers A^k and A^l $(0 \leq k < l \leq j)$ lie in the same coset, and A^g, with $g = l - k$, is an element of G_1, q. e. d.

53. LEMMA 22. *Let Δ^* be a subgroup of $\Delta_0(r, s)$ of finite index. Then there exists in Δ^* a diagonal matrix $\mathfrak{P}$, such that all diagonal elements of $\mathfrak{P}$ are different one from another and that no conjugate of $\mathfrak{P}$, different from $\mathfrak{P}$ and $\mathfrak{P}^{-1}$, has a diagonal element in common with $\mathfrak{P}$.*

Let h be the degree of the field K with the conjugates $K^{(1)}, \cdots, K^{(h)}$ and $K^{(1)} = K$. The number of algebraic fundamental units in K is $h - 1$. The field $K(\sqrt{rs})$ has the degree $2h$; since it has 2 real conjugates and $h - 1$ pairs of conjugate complex conjugates, the number of algebraic fundamental units in $K(\sqrt{rs})$ is h. Consequently there exists an algebraic unit $\lambda = a + b\sqrt{rs}$, where a and b are numbers of K, such that no power λ^q $(q = \pm 1, \pm 2, \cdots)$ is a number of K.

Denoting by N the norm relative to K, we have $N(\lambda) = a^2 - rsb^2 = c$, where c is an algebraic unit in K, and $N(c^{-1}\lambda^2) = 1$. We replace $c^{-1}\lambda^2$ again by λ; then λ is an algebraic unit in $K(\sqrt{rs})$, no power λ^q $(q = \pm 1, \pm 2, \cdots)$ is a number of K, and $N(\lambda) = 1$, $\lambda^{-1} = a - b\sqrt{rs}$. By Fermat's theorem, we may, moreover, assume $\lambda \equiv 1 \pmod{2\sqrt{rs}}$; hence a and b are integers of K.

Let $q_1, \cdots, q_n$ be different positive rational integers and let

$$\mathfrak{P}_1 = [\lambda^{q_1}, \cdots, \lambda^{q_n}].$$

Then the matrix

$$\mathfrak{P} = \begin{pmatrix} \mathfrak{P}_1 & 0 \\ 0 & \mathfrak{P}_1^{-1} \end{pmatrix} = \tfrac{1}{2}(\mathfrak{P}_1 + \mathfrak{P}_1^{-1}) \times \epsilon_0 + \frac{1}{2\sqrt{rs}}(\mathfrak{P}_1 - \mathfrak{P}_1^{-1}) \times \epsilon_1$$

is an element of the group $\Delta_0(r,s)$. By Lemma 21, a certain power $\mathfrak{P}^g$ with positive exponent g lies in the subgroup Δ^*. Replacing gq_k $(k=1,\cdots,n)$ by q_k, we may assume that already $\mathfrak{P}$ is an element of Δ^*. The diagonal elements of $\mathfrak{P}$ are $\lambda^{q_k}, \lambda^{-q_k}$ $(k=1,\cdots,n)$. Since the exponents $q_k, -q_k$ are all different and λ is no root of unity, these diagonal elements are all different.

For any number t of K, we denote by $t^{(l)}$ the conjugate of t in $K^{(l)}$. The $2h$ conjugates of $\lambda^{q_k} = a_k + b_k\sqrt{rs}$ are $a_k^{(l)} \pm b_k^{(l)}\sqrt{r^{(l)}s^{(l)}}$ $(l=1,\cdots,h)$. Since $b_k \neq 0$ and $r^{(l)}s^{(l)} < 0$ for $l=2,\cdots,h$, only the two conjugates $a_k + b_k\sqrt{rs} = \lambda^{q_k}$ and $a_k - b_k\sqrt{rs} = \lambda^{-q_k}$ are real. If a conjugate $\mathfrak{P}^*$ of $\mathfrak{P}$ has a diagonal element in common with $\mathfrak{P}$, then the substitution $\mathfrak{P} \to \mathfrak{P}^*$ arises either from the identical mapping $\lambda \to \lambda$ or from $\lambda \to \lambda^{-1}$, hence $\mathfrak{P}^* = \mathfrak{P}$ or $\mathfrak{P}^* = \mathfrak{P}^{-1}$, q. e. d.

We assume that the groups $\Delta_0(r,s)$ and $\Delta_0(r_1,s_1)$ are commensurable; we shall now first prove, that then $K = K_1$. There exists a subgroup Δ^* of $\Delta_0(r,s)$, of finite index, and a symplectic matrix $\mathfrak{R}$, such that $\mathfrak{R}^{-1}\Delta^*\mathfrak{R} = \Delta_1$ is a subgroup of $\Delta_0(r_1,s_1)$, of finite index. Henceforth we do not need the existence of a *symplectic* matrix $\mathfrak{R}$ with this property; we have only to assume that there is a non-singular matrix $\mathfrak{R}$ with real or complex elements satisfying $\Delta^*\mathfrak{R} = \mathfrak{R}\Delta_1$ where Δ^* and Δ_1 are subgroups of $\Delta_0(r,s)$ and $\Delta_0(r_1,s_1)$, of finite indices. Obviously we may then, moreover, assume that $\mathfrak{R}$ lies in the field $K^*(\sqrt{r}, \sqrt{s}, \sqrt{r_1}, \sqrt{s_1}) = K_0$, of degree 16 over the union K^* of K and K_1.

Let $\mathfrak{P}$ be the matrix of Lemma 22. Since $\mathfrak{P}$ belongs to the subgroup Δ^* of $\Delta_0(r,s)$, the matrix $\mathfrak{R}^{-1}\mathfrak{P}\mathfrak{R}$ is an element of Δ_1 and lies therefore in the field $K_1(\sqrt{r_1}, \sqrt{s_1})$. We consider any isomorphism A of K_0 which does not change the numbers of $K_1(\sqrt{r_1}, \sqrt{s_1})$. Denoting the image under A by the subscript A, we have $\mathfrak{R}_A^{-1}\mathfrak{P}_A\mathfrak{R}_A = \mathfrak{R}^{-1}\mathfrak{P}\mathfrak{R}$; hence the matrix $\mathfrak{V} = \mathfrak{R}_A\mathfrak{R}^{-1}$ satisfies

$$\mathfrak{P}_A = \mathfrak{V}\mathfrak{P}\mathfrak{V}^{-1}. \tag{155}$$

This proves that the diagonal matrices $\mathfrak{P}_A$ and $\mathfrak{P}$ have the same diagonal elements, perhaps in different order. By Lemma 22, either $\mathfrak{P}_A = \mathfrak{P}$ or $\mathfrak{P}_A = \mathfrak{P}^{-1}$. If $a + b\sqrt{rs}$ is a diagonal element of $\mathfrak{P}$, then $b \neq 0$ and $(a + b\sqrt{rs})_A = a \pm b\sqrt{rs}$; whence $(rs)_A = rs$.

On the other hand, rs generates K, since all conjugates of rs except rs itself are negative. Consequently all numbers of K are invariant under A. This proves $K \subset K_1(\sqrt{r_1}, \sqrt{s_1})$. Since $K^*(\sqrt{r_1}, \sqrt{s_1})$ has the degree 4 relative to K^*, the intersection of $K_1(\sqrt{r_1}, \sqrt{s_1})$ and K^* is K_1. Therefore

$K \subset K_1$. Interchanging K and K_1, we have also $K_1 \subset K$; consequently $K = K_1$.

54. We use the abbreviation

$$\mathfrak{C} = \begin{pmatrix} \mathfrak{E} & \mathfrak{E} \\ -\mathfrak{E} & \mathfrak{E} \end{pmatrix}.$$

LEMMA 23. *There exists an element $\mathfrak{M}$ of Δ^*, such that $\mathfrak{C}\mathfrak{M}\mathfrak{C}^{-1} = \mathfrak{Q}$ is a diagonal matrix with different diagonal elements.*

Analogously to the proof of Lemma 22, we construct an algebraic unit $\mu = a + b\sqrt{s}$ in $K(\sqrt{s})$, such that a, b are integers of K, $a^2 - sb^2 = 1$ and all powers μ^q $(q = \pm 1, \pm 2, \cdots)$ are different one from another. Let $q_1, \cdots, q_n$ be different positive rational integers, $\mathfrak{Q}_1 = [\mu^{q_1}, \cdots \mu^{q_n}]$ and

$$\mathfrak{Q} = \begin{pmatrix} \mathfrak{Q}_1 & 0 \\ 0 & \mathfrak{Q}_1^{-1} \end{pmatrix}.$$

Then the matrix

$$\mathfrak{M} = \mathfrak{C}^{-1}\mathfrak{Q}\mathfrak{C} = \tfrac{1}{2}\begin{pmatrix} \mathfrak{Q}_1 + \mathfrak{Q}_1^{-1} & \mathfrak{Q}_1 - \mathfrak{Q}_1^{-1} \\ \mathfrak{Q}_1 - \mathfrak{Q}_1^{-1} & \mathfrak{Q}_1 + \mathfrak{Q}_1^{-1} \end{pmatrix}$$

$$= \tfrac{1}{2}(\mathfrak{Q}_1 + \mathfrak{Q}_1^{-1}) \times \epsilon_0 + \frac{1}{2\sqrt{s}}(\mathfrak{Q}_1 - \mathfrak{Q}_1^{-1}) \times \epsilon_3$$

is an element of $\Delta_0(r, s)$. By Lemma 3, a power $\mathfrak{M}^g = \mathfrak{C}^{-1}\mathfrak{Q}^g\mathfrak{C}$ lies in Δ^*. Replacing $\mathfrak{M}^g, \mathfrak{Q}^g$ by $\mathfrak{M}, \mathfrak{Q}$, we obtain the proof of Lemma 23.

Let again A denote an isomorphism of $K(\sqrt{r}, \sqrt{s}, \sqrt{r_1}, \sqrt{s_1}) = K_0$ which leaves invariant all numbers of $K(\sqrt{r_1}, \sqrt{s_1})$. Analogous to (155), we obtain

(156) $$\mathfrak{Q}_A = (\mathfrak{C}\mathfrak{B}\mathfrak{C}^{-1})\mathfrak{Q}(\mathfrak{C}\mathfrak{B}\mathfrak{C}^{-1})^{-1}$$

with the matrix $\mathfrak{Q}$ of Lemma 23 and $\mathfrak{B} = \mathfrak{R}_A\mathfrak{R}^{-1}$. Consider in particular the isomorphism A_s defined by $\sqrt{s} \to -\sqrt{s}$, $\sqrt{r} \to \sqrt{r}$ and write more shortly the subscript s instead of A_s; then

$$\mathfrak{P}_s = \mathfrak{P}^{-1}, \qquad \mathfrak{Q}_s = \mathfrak{Q}^{-1}.$$

By (155) the matrix $\mathfrak{R}_s\mathfrak{R}^{-1} = \mathfrak{B}$ has the form

$$\mathfrak{B} = \begin{pmatrix} 0 & \mathfrak{D}(1) \\ \mathfrak{D}(0) & 0 \end{pmatrix},$$

where $\mathfrak{D}(0)$ and $\mathfrak{D}(1)$ denote diagonal matrices; by (156), the matrix $\mathfrak{C}\mathfrak{B}\mathfrak{C}^{-1}$ has the same form, whence

$$\mathfrak{R}_s = \begin{pmatrix} 0 & \mathfrak{D}(1) \\ -\mathfrak{D}(1) & 0 \end{pmatrix} \mathfrak{R}, \tag{157}$$

$$\mathfrak{R} = \mathfrak{R}_{ss} = -(\mathfrak{D}_s(1)\mathfrak{D}(1) \times \epsilon_0)\mathfrak{R}, \qquad \mathfrak{D}_s(1)\mathfrak{D}(1) = -\mathfrak{E}. \tag{158}$$

Consider now the isomorphism A_r defined by $\sqrt{r} \to -\sqrt{r}$, $\sqrt{s} \to \sqrt{s}$; then

$$\mathfrak{P}_r = \mathfrak{P}^{-1}, \qquad \mathfrak{Q}_r = \mathfrak{Q},$$

and we obtain from (155) and (156) the formula

$$\mathfrak{R}_r = \begin{pmatrix} 0 & \mathfrak{D}(2) \\ \mathfrak{D}(2) & 0 \end{pmatrix} \mathfrak{R} \tag{159}$$

with a diagonal matrix $\mathfrak{D}(2)$, whence

$$\mathfrak{D}_r(2)\mathfrak{D}(2) = \mathfrak{E}. \tag{160}$$

Interchanging r, s and r_1, s_1, we find in an analogous manner

$$\mathfrak{R}_{s_1} = \mathfrak{R}\begin{pmatrix} 0 & \mathfrak{D}(3) \\ -\mathfrak{D}(3) & 0 \end{pmatrix}, \qquad \mathfrak{R}_{r_1} = \mathfrak{R}\begin{pmatrix} 0 & \mathfrak{D}(4) \\ \mathfrak{D}(4) & 0 \end{pmatrix}, \tag{161}$$

$$\mathfrak{D}(3)\mathfrak{D}_{s_1}(3) = -\mathfrak{E}, \qquad \mathfrak{D}(4)\mathfrak{D}_{r_1}(4) = \mathfrak{E} \tag{162}$$

with diagonal matrices $\mathfrak{D}(3)$ and $\mathfrak{D}(4)$. Moreover $\mathfrak{R}_{rr_1} = \mathfrak{R}_{r_1 r}$, whence

$$(\mathfrak{D}_{r_1}(2)\mathfrak{D}^{-1}(2) \times \epsilon_0)\mathfrak{R} = \mathfrak{R}(\mathfrak{D}_r(4)\mathfrak{D}^{-1}(4) \times \epsilon_0). \tag{163}$$

Since $|\mathfrak{R}| \neq 0$ at least one element r_{1l} of the first row of $\mathfrak{R} = (r_{kl})$ is $\neq 0$. We assume $r_{1l} \neq 0$ for $l = p_0$ and define $p = p_0$, if $p_0 \leq n$, and $p = p_0 - n$, if $p > n$. Let a be the first diagonal element of $\mathfrak{D}(2)$ and b be the p-th diagonal element of $\mathfrak{D}(4)$. By (160), (162) and (163),

$$aa_r = 1, \qquad bb_{r_1} = 1, \qquad ab_r = ba_{r_1}.$$

55. LEMMA 24. *There exists a number $c \neq 0$ in K_0 satisfying*

$$ac = c_r, \qquad bc = c_{r_1}. \tag{164}$$

If $a \neq 1$, we define $d = \dfrac{\sqrt{r}}{a-1}$, whence

$$d_r = \frac{-\sqrt{r}}{a_r - 1} = \frac{-a\sqrt{r}}{1-a} = ad;$$

if $a = 1$, we define $d = 1$. In both cases $ad = d_r$. Putting $\dfrac{bd}{d_{r_1}} = f$, we find $ff_{r_1} = 1$,

$$f_r = \frac{b_r d_r}{d_{r_1 r}} = \frac{b_r a d}{(ad)_{r_1}} = \frac{bd}{d_{r_1}} = f.$$

Let $g = \dfrac{\sqrt{r_1}}{f-1}$ for $f \neq 1$, and $g = 1$ for $f = 1$. In both cases

$$fg = g_{r_1}, \qquad g_r = g,$$

and $c = dg$ has the required propetries.

Since there is an arbitrary scalar factor $\neq 0$ in $\mathfrak{R}$, we may replace $\mathfrak{R}$ by $c\mathfrak{R}$, where c is defined in Lemma 24. It follows from (159), (161) and (164), that then a and b are both replaced by 1.

Let α be the first diagonal element of $\mathfrak{D}(1)$ and β the p-th diagonal element of $\mathfrak{D}(3)$. By (158) and (162),

$$\alpha\alpha_s = -1, \qquad \beta\beta_{s_1} = -1. \tag{165}$$

Calculating

$$\mathfrak{R}_{rs} = \mathfrak{R}_{sr}, \quad \mathfrak{R}_{sr_1} = \mathfrak{R}_{r_1 s}, \quad \mathfrak{R}_{rs_1} = \mathfrak{R}_{s_1 r}, \quad \mathfrak{R}_{r_1 s_1} = \mathfrak{R}_{s_1 r_1} \text{ and } \quad \mathfrak{R}_{ss_1} = \mathfrak{R}_{s_1 s}$$

by (157), (159) and (161), we obtain, moreover,

$$\alpha_r = -\alpha, \quad \alpha_{r_1} = \alpha, \quad \beta_r = \beta, \quad \beta_{r_1} = -\beta, \quad \beta\alpha_{s_1} = \alpha\beta_s. \tag{166}$$

By (165) and (166), the numbers $u = \alpha\sqrt{r}$ and $v = \beta\sqrt{r_1}$ lie in $K(\sqrt{s}, \sqrt{s_1})$ and satisfy

$$uu_s = -r, \qquad vv_{s_1} = -r_1, \qquad vu_{s_1} = uv_s.$$

Defining $w = u/u_{s_1} = v/v_s$, we have $w_{s_1} = w^{-1} = w_s$, and consequently w is a number of the field $K(\sqrt{ss_1})$. Let $\tau = \sqrt{ss_1}/(w-1)$ for $w \neq 1$, and $\tau = 1$ for $w = 1$; in both cases τ lies in $K(\sqrt{ss_1})$ and $\tau_s = \tau_{s_1} = w\tau$. Then the numbers $\rho = u\tau$ and $\sigma = v\tau$ satisfy

$$\rho_{s_1} = u_{s_1}\tau_{s_1} = \frac{u}{w}w\tau = u\tau = \rho, \qquad \sigma_s = v_s\tau_s = \frac{v}{w}w\tau = v\tau = \sigma,$$

$$\rho\rho_s = u\tau u_s\tau_s = -r\tau\tau_s, \qquad \sigma\sigma_{s_1} = v\tau v_{s_1}\tau_{s_1} = -r_1\tau\tau_{s_1},$$

whence

$$\rho = \xi + \eta\sqrt{s}, \quad \sigma = \xi_1 + \eta_1\sqrt{s_1}, \quad \tau = \zeta + \omega\sqrt{ss_1},$$

$$\xi^2 - s\eta^2 = -r(\zeta^2 - ss_1\omega^2), \quad \xi_1^2 - s_1\eta_1^2 = -r_1(\zeta^2 - ss_1\omega^2), \tag{167}$$

where ξ, η, ξ_1, η_1, ζ, ω are numbers of K which are not all 0.

We consider first the case $\omega \neq 0$. Performing the 3 linear substitutions

$$x = \eta x_3 + \xi y_3, \qquad y = \xi x_3 + s\eta y_3, \qquad z = z_3$$

$$x_3 = \frac{x_2}{s\eta^2 - \xi^2}, \qquad y_3 = \frac{\zeta y_2}{s\omega(s\eta^2 - \xi^2)} + \frac{z_2}{rs\omega}, \qquad z_3 = \frac{y_2 + \zeta z_2}{s\omega}$$

$$x_2 = s_1\eta_1 x_1 + \xi_1 y_1, \qquad y_2 = \xi_1 x_1 + \eta_1 y_1, \qquad z_2 = z_1,$$

we obtain, by (167),

$$Q(r,s) = rsx^2 - ry^2 + sz^2 = r(s\eta^2 - \xi^2)(x_3^2 - sy_3^2) + sz_3^2$$
$$= \frac{r}{s\eta^2 - \xi^2}(x_2^2 - s_1 y_2^2) + s_1 z_2^2 = r_1 s_1 x_1^2 - r_1 y_1^2 + s_1 z_1^2 = Q(r_1, s_1);$$

hence $Q(r,s)$ and $Q(r_1,s_1)$ are equivalent in K.

In the remaining case $\omega = 0$, we have

$$\xi^2 - s\eta^2 = -r\zeta^2 \quad \text{and} \quad \xi_1^2 - s_1\eta_1^2 = -r_1\zeta^2.$$

Consequently the diophantine equation $Q(r,s) = 0$ has the non-trivial solution $x = \xi$, $y = s\eta$, $z = r\zeta$. Moreover it follows immediately from the signs of the conjugates of r and s, that K is the field of rational numbers. Applying the 2 linear substitutions

$$x = \frac{\eta x_2 + \xi y_2}{r\zeta}, \qquad y = \frac{\xi x_2 + s\eta y_2}{r\zeta}, \qquad z = z_2$$

$$x_2 = x_1, \qquad y_2 = \frac{s+1}{2s} y_1 + \frac{s-1}{2s} z_1, \qquad z_2 = \frac{s-1}{2s} y_1 + \frac{s+1}{2s} z_1,$$

we obtain

$$Q(r,s) = x_2^2 - sy_2^2 + sz_2^2 = x_1^2 - y_1^2 + z_1^2 = Q(1,1);$$

hence $Q(r,s)$ and $Q(1,1)$ are equivalent in K. Since the same holds for $Q(r_1,s_1)$ instead of $Q(r,s)$, $Q(r,s)$ and $Q(r_1,s_1)$ are also equivalent in K.

Theorem 13 is now completely proved. In the particular case

$$\mathfrak{G}_1 = \mathfrak{J}, \quad \mathfrak{H}_1 = i\mathfrak{J}, \quad r_1 = 1, \quad s_1 = 1,$$

we have

$$\Delta(\mathfrak{G}_1, \mathfrak{H}_1) = \Gamma \quad \text{and} \quad Q(r_1,s_1) = x_1^2 - y_1^2 + z_1^2.$$

It follows that the group $\Delta(\mathfrak{G}, \mathfrak{H})$ is then and only then commensurable with the modular group, when the diophantine equation $x^2 + ry^2 = s$ has a solution x, y in K.

X. UNIT GROUPS OF QUINARY QUADRATIC FORMS.

56. For $n > 2$, the groups $\Delta(\mathfrak{G}, \mathfrak{H})$ and their subgroups are the only known examples of non-trivial discontinuous subgroups of the symplectic group. However, in the case $n = 2$, another set of examples is provided by the unit groups of certain quinary quadratic forms.

Consider the special quinary quadratic form

$$\mathfrak{S}[\mathfrak{w}] = w_1 w_2 - w_3^2 + w_4 w_5$$

and a complex column $\mathfrak{w}$ satisfying

$$\mathfrak{S}[\mathfrak{w}] = 0, \qquad \mathfrak{S}\{\mathfrak{w}\} > 0. \tag{168}$$

If $w_5 = 0$, then $w_1 w_2 = w_3^2$ and $w_1 \bar{w}_2 + w_2 \bar{w}_1 - 2 w_3 \bar{w}_3 > 0$, whence $w_1 \neq 0$,

$$(w_1 \bar{w}_3 - w_3 \bar{w}_1)^2 = w_1 \bar{w}_1 (w_1 \bar{w}_2 + w_2 \bar{w}_1 - 2 w_3 \bar{w}_3) > 0,$$

which is impossible, the left-hand side being the square of a pure imaginary quantity; consequently $w_5 \neq 0$. Introducing inhomogeneous coördinates $w_5^{-1}\mathfrak{w} = \mathfrak{z} = \mathfrak{x} + i\mathfrak{y}$, we infer from

$$\mathfrak{S}[\mathfrak{z}] = 0 \quad \text{and} \quad \mathfrak{S}[\mathfrak{y}] = \mathfrak{S}\left[\frac{\mathfrak{z} - \bar{\mathfrak{z}}}{2i}\right] = \tfrac{1}{2}\mathfrak{S}\{\mathfrak{z}\} > 0$$

the relationships

$$z_1 z_2 - z_3^2 = z_4, \qquad y_1 y_2 - y_3^2 > 0, \tag{169}$$

whence, in particular, $y_1 \neq 0$. On the other hand, (168) follows again from (169), if we define $\mathfrak{w} = w_5 \mathfrak{z}$ with arbitrary $w_5 \neq 0$.

The matrices

$$\mathfrak{Z} = \begin{pmatrix} z_1 & z_3 \\ z_3 & z_2 \end{pmatrix}$$

with $y_1 y_2 - y_3^2 > 0$ and $y_1 > 0$ form the space H for $n = 2$; let $\bar{H}$ be the space defined by $y_1 y_2 - y_3^2 > 0$ and $y_1 < 0$.

Consider now a real linear transformation $\hat{\mathfrak{w}} = \mathfrak{V}\mathfrak{w}$, such that $\mathfrak{S}[\hat{\mathfrak{w}}] = 0$, $\mathfrak{S}\{\hat{\mathfrak{w}}\} > 0$ follows from (168). Since the equation $\mathfrak{S}[\mathfrak{w}] = 0$ is irreducible, we conclude that $\mathfrak{S}[\mathfrak{V}] = \lambda \mathfrak{S}$, with a scalar factor $\lambda \neq 0$. Consequently $|\mathfrak{V}| \neq 0$, and these transformations form a group. Since $w_1, \cdots, w_5$ are homogeneous coördinates, we may replace $\mathfrak{V}$ by $\mu\mathfrak{V}$, for any scalar $\mu \neq 0$. But $|\mu\mathfrak{V}| = \mu^5 |\mathfrak{V}|$; hence we may assume $|\mathfrak{V}| = 1$ and $\mathfrak{S}[\mathfrak{V}] = \mathfrak{S}$.

Put $\mathfrak{V} = (v_{kl})$ and $\hat{w}_5^{-1}\hat{\mathfrak{w}} = \hat{\mathfrak{z}} = \hat{\mathfrak{x}} + i\hat{\mathfrak{y}}$. Then the fractional linear substitution

$$\hat{z}_k = \frac{\sum_{l=1}^{5} v_{kl} z_l}{\sum_{l=1}^{5} v_{5l} z_l} \qquad (k = 1, 2, 3) \tag{170}$$

with $z_4 = z_1 z_2 - z_3^2$ and $z_5 = 1$ transforms the space H either into itself or into $\bar{H}$. We are only interested in the substitutions which leave H invariant; they form an invariant subgroup of index 2 in the whole group. We denote this subgroup by $\Omega(\mathfrak{S})$.

Obviously we obtain the solution of (170) with respect to z_1, z_2, z_3, if we replace $\mathfrak{V}$ by $\mathfrak{V}^{-1}$ and interchange $\mathfrak{z}$, $\hat{\mathfrak{z}}$. Hence (170) is a birational analytic mapping of H onto itself. A simple calculation gives the functional determinant

$$\frac{d(\hat{z}_1, \hat{z}_2, \hat{z}_3)}{d(z_1, z_2, z_3)} = \hat{w}_5^{-3} w_5^{3} \neq 0. \tag{171}$$

On account of Theorem 1, the transformation (170) is symplectic:

$$\begin{pmatrix} \hat{z}_1 & \hat{z}_3 \\ \hat{z}_3 & \hat{z}_2 \end{pmatrix} = \hat{\mathfrak{Z}} = (\mathfrak{A}\mathfrak{Z} + \mathfrak{B})(\mathfrak{C}\mathfrak{Z} + \mathfrak{D})^{-1}, \quad \begin{pmatrix} \mathfrak{A} & \mathfrak{B} \\ \mathfrak{C} & \mathfrak{D} \end{pmatrix} = \mathfrak{M}, \quad \mathfrak{J}[\mathfrak{M}] = \mathfrak{J}. \tag{172}$$

57. Let us now start from an arbitrary symplectic transformation (172). Defining

$$\mathfrak{J}_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

we have $\mathfrak{J}_2[\mathfrak{F}] = |\mathfrak{F}|\mathfrak{J}_2$ for any matrix $\mathfrak{F}^{(2)}$, and consequently

$$(\mathfrak{C}\mathfrak{Z} + \mathfrak{D})^{-1} = |\mathfrak{C}\mathfrak{Z} + \mathfrak{D}|^{-1}\mathfrak{J}_2(\mathfrak{Z}\mathfrak{C}' + \mathfrak{D}')\mathfrak{J}_2^{-1}. \tag{173}$$

Moreover

$$|\mathfrak{F}_1 + \mathfrak{F}_2| = |\mathfrak{F}_1| + |\mathfrak{F}_2| + \sigma(\mathfrak{F}_1\mathfrak{J}_2\mathfrak{F}'_2\mathfrak{J}_2^{-1})$$

for any two matrices $\mathfrak{F}_1^{(2)}$ and $\mathfrak{F}_2^{(2)}$, whence

$$|\mathfrak{A}\mathfrak{Z} + \mathfrak{B}| = |\mathfrak{A}\mathfrak{Z}| + |\mathfrak{B}| + \sigma(\mathfrak{A}\mathfrak{Z}\mathfrak{J}_2\mathfrak{B}'\mathfrak{J}_2^{-1}) \tag{174}$$

$$|\mathfrak{C}\mathfrak{Z} + \mathfrak{D}| = |\mathfrak{C}\mathfrak{Z}| + |\mathfrak{D}| + \sigma(\mathfrak{C}\mathfrak{Z}\mathfrak{J}_2\mathfrak{D}'\mathfrak{J}_2^{-1}). \tag{175}$$

We apply (173), (174) and (175) for the calculation of

$$\hat{\mathfrak{Z}} = (\mathfrak{A}\mathfrak{Z} + \mathfrak{B})(\mathfrak{C}\mathfrak{Z} + \mathfrak{D})^{-1} \quad \text{and} \quad |\hat{\mathfrak{Z}}| = |\mathfrak{A}\mathfrak{Z} + \mathfrak{B}|\,|\mathfrak{C}\mathfrak{Z} + \mathfrak{D}|^{-1}.$$

It follows that

$$\hat{\mathfrak{W}} = (\mathfrak{A}\mathfrak{W}\mathfrak{J}_2\mathfrak{D}' + \mathfrak{B}\mathfrak{J}_2\mathfrak{W}\mathfrak{C}' + w_4\mathfrak{A}\mathfrak{J}_2\mathfrak{C}' + w_5\mathfrak{B}\mathfrak{J}_2\mathfrak{D}')\mathfrak{J}_2^{-1} \tag{176}$$

$$\hat{w}_4 = \sigma(\mathfrak{A}\mathfrak{W}\mathfrak{J}_2\mathfrak{B}'\mathfrak{J}_2^{-1}) + |\mathfrak{A}|w_4 + |\mathfrak{B}|w_5 \tag{177}$$

$$\hat{w}_5 = \sigma(\mathfrak{C}\mathfrak{W}\mathfrak{J}_2\mathfrak{D}'\mathfrak{J}_2^{-1}) + |\mathfrak{C}|w_4 + |\mathfrak{D}|w_5 \tag{178}$$

with

$$\mathfrak{W} = \begin{pmatrix} w_1 & w_3 \\ w_3 & w_2 \end{pmatrix}, \qquad \hat{\mathfrak{W}} = \begin{pmatrix} \hat{w}_1 & \hat{w}_3 \\ \hat{w}_3 & \hat{w}_2 \end{pmatrix}$$

is a real linear substitution $\hat{\mathfrak{w}} = \mathfrak{V}\mathfrak{w}$ mapping the domain $\mathfrak{S}[\mathfrak{w}] = 0$, $\mathfrak{S}\{\mathfrak{w}\} > 0$ into itself. Hence $|\mathfrak{B}| = \nu \neq 0$, $|\nu^{-1/5}\mathfrak{B}| = 1$. Moreover, $\mathfrak{B}$ is not changed if we replace $\mathfrak{M}$ my $-\mathfrak{M}$. This proves the identity of $\Omega(\mathfrak{S})$ and the symplectic group Ω, for $n = 2$. By (171) and (178), the functional determinant of the transformation (172) has the value

$$(\nu^{-1/5}\hat{w}_5)^{-3}w_5^{\,3} = \nu^{3/5}\,|\,\mathfrak{C}\mathfrak{Z} + \mathfrak{D}\,|^{-3};$$

on the other hand, by (24), $d\mathfrak{Z} = d\hat{\mathfrak{Z}}[\mathfrak{C}\mathfrak{Z} + \mathfrak{D}]$, which leads to the value $|\mathfrak{C}\mathfrak{Z} + \mathfrak{D}|^{-3}$ of that functional determinant; consequently $\nu = 1$, $|\mathfrak{V}| = 1$ and $\mathfrak{S}[\mathfrak{V}] = \mathfrak{S}$.

58. The formulae (176), (177), (178) and the identity of the groups $\Omega(\mathfrak{S})$ and Ω can be demonstrated in another way, without using Theorem 1.

Let us first determine all skew-symmetric symplectic matrices $\mathfrak{G}^{(4)}$. We have the conditions

$$\mathfrak{G} = \begin{pmatrix} \mathfrak{A} & \mathfrak{B} \\ \mathfrak{C} & \mathfrak{D} \end{pmatrix} = -\mathfrak{G}' \tag{179}$$

$$\mathfrak{A}\mathfrak{B}' = \mathfrak{B}\mathfrak{A}', \qquad \mathfrak{C}\mathfrak{D}' = \mathfrak{D}\mathfrak{C}', \qquad \mathfrak{A}\mathfrak{D}' - \mathfrak{B}\mathfrak{C}' = \mathfrak{E}. \tag{180}$$

By (179),

$$\mathfrak{A} = w_4\mathfrak{J}_2, \quad \mathfrak{D} = w_5\mathfrak{J}_2, \quad \mathfrak{B} = \mathfrak{W}\mathfrak{J}_2, \quad \mathfrak{C} = \mathfrak{J}_2\mathfrak{W}', \quad \mathfrak{W} = \begin{pmatrix} w_1 & w_3 \\ w_6 & w_2 \end{pmatrix}$$

with arbitrary $w_1, \cdots, w_6$; by (180), $w_4\mathfrak{W}$ and $w_5\mathfrak{W}$ are symmetric and $\mathfrak{W}\mathfrak{J}_2\mathfrak{W} = (w_4w_5 - 1)\mathfrak{J}_2$. Hence either $\mathfrak{W}$ is non-symmetric and then, necessarily, $\mathfrak{G} = \pm\mathfrak{J}$ or $\mathfrak{W}$ is symmetric and

$$w_1w_2 - w_3^{\,2} = w_4w_5 - 1. \tag{181}$$

Omitting in the second case the condition (181), we obtain the identity

$$(\mathfrak{J}\mathfrak{G})^2 = -\mathfrak{S}[\mathfrak{w}]\mathfrak{E} \tag{182}$$

with

$$\mathfrak{G} = \begin{pmatrix} w_4\mathfrak{J}_2 & \mathfrak{W}\mathfrak{J}_2 \\ \mathfrak{J}_2\mathfrak{W} & w_5\mathfrak{J}_2 \end{pmatrix}, \qquad \mathfrak{W} = \begin{pmatrix} w_1 & w_3 \\ w_3 & w_2 \end{pmatrix}$$

and arbitrary $w_1, \cdots, w_5$.

We substitute

$$\begin{aligned} &w_1 = -v_1 - v_2, \quad w_2 = v_1 - v_2, \quad w_3 = -v_3, \\ &w_4 = -v_4 - v_5, \quad w_5 = v_4 - v_5; \end{aligned} \tag{183}$$

then

$$\begin{aligned} \mathfrak{J}\mathfrak{G} &= \begin{pmatrix} \mathfrak{J}_2\mathfrak{W} & w_5\mathfrak{J}_2 \\ -w_4\mathfrak{J}_2 & -\mathfrak{W}\mathfrak{J}_2 \end{pmatrix} \\ &= \begin{pmatrix} -v_3 & v_1 - v_2 & 0 & v_4 - v_5 \\ v_1 + v_2 & v_3 & -v_4 + v_5 & 0 \\ 0 & v_4 + v_5 & -v_3 & v_1 + v_2 \\ -v_4 - v_5 & 0 & v_1 - v_2 & v_3 \end{pmatrix} = \sum_{k=1}^{5} v_k\mathfrak{F}_k, \end{aligned} \tag{184}$$

$$\mathfrak{S}[\mathfrak{w}] = \sum_{k=1}^{5} (-1)^k v_k^2,$$

$$\begin{aligned} \mathfrak{F}_k^2 &= (-1)^{k-1}\mathfrak{E} && (k = 1, \cdots, 5), \\ \mathfrak{F}_k\mathfrak{F}_l &= -\mathfrak{F}_l\mathfrak{F}_k && (1 \leq k < l \leq 5), \end{aligned} \tag{185}$$

by (182). The 16 products $\mathfrak{F}_1^{e_1}\mathfrak{F}_2^{e_2}\mathfrak{F}_3^{e_3}\mathfrak{F}_4^{e_4}$ $(e_k = 0, 1;\ k = 1, \cdots, 4)$ form a basis for the four-dimensional complete matric representation of the well-known Clifford-Lipschitz algebra of order 16. The matrix $\mathfrak{F}_1\mathfrak{F}_2\mathfrak{F}_3\mathfrak{F}_4\mathfrak{F}_5$ is permutable with all matrices of the algebra and has the square $\mathfrak{E}$, hence it is equal to $\mathfrak{E}$ or $-\mathfrak{E}$; by direct calculation we find the value $\mathfrak{E}$ and therefore

$$\mathfrak{F}_5 = \mathfrak{F}_1\mathfrak{F}_2\mathfrak{F}_3\mathfrak{F}_4. \tag{186}$$

Let now $\hat{\mathfrak{w}} = \mathfrak{B}\mathfrak{w}$ be a real linear substitution with $\mathfrak{S}[\hat{\mathfrak{w}}] = \mathfrak{S}[\mathfrak{w}]$ and $|\mathfrak{B}| = 1$. We denote the matrix of the substitution (183) by $\mathfrak{L}$ and put $\mathfrak{L}^{-1}\mathfrak{B}\mathfrak{L} = \mathfrak{R} = (r_{kl})$. By (182) and (184), the matrices

$$\hat{\mathfrak{F}}_k = \sum_{l=1}^{5} r_{lk}\mathfrak{F}_l \qquad (k = 1, \cdots, 5) \tag{187}$$

satisfy again (185), and consequently their product is $\pm\mathfrak{E}$. It follows from (185), (186) and the linear independence of the 16 products that

$$\hat{\mathfrak{F}}_1\hat{\mathfrak{F}}_2\hat{\mathfrak{F}}_3\hat{\mathfrak{F}}_4\hat{\mathfrak{F}}_5 = |\mathfrak{R}|\,\mathfrak{E} = \mathfrak{E}; \tag{188}$$

hence (186) holds also for $\hat{\mathfrak{F}}_k$ instead of $\mathfrak{F}_k$.

Since the two representations generated by $\mathfrak{F}_k$ and $\hat{\mathfrak{F}}_k$ are necessarily equivalent, there exists a real matrix $\mathfrak{M}$, such that

$$\hat{\mathfrak{F}}_k = \mathfrak{M}^{-1}\mathfrak{F}_k\mathfrak{M} \qquad (k = 1, \cdots, 5). \tag{189}$$

Moreover, by (184) and (187), the matrices $\mathfrak{J}\mathfrak{F}_k$ and $\mathfrak{J}\hat{\mathfrak{F}}_k$ are skew-symmetric, and consequently (189) leads to

$$\mathfrak{J}\mathfrak{M}^{-1}\mathfrak{F}_k\mathfrak{M} = \mathfrak{M}'\mathfrak{J}\mathfrak{F}_k\mathfrak{J}\mathfrak{M}'^{-1}\mathfrak{J}^{-1} \qquad (k = 1, \cdots, 5).$$

This proves that $\mathfrak{M}\mathfrak{J}\mathfrak{M}'\mathfrak{J}^{-1}$ is a scalar multiple of $\mathfrak{E}$. There is an arbitrary real scalar factor in $\mathfrak{M}$; hence we may assume $\mathfrak{J}[\mathfrak{M}] = \pm\mathfrak{J}$. Putting

$$\hat{\mathfrak{G}} = \begin{pmatrix} \hat{w}_4\mathfrak{J}_2 & \hat{\mathfrak{W}}\mathfrak{J}_2 \\ \mathfrak{J}_2\hat{\mathfrak{W}} & \hat{w}_5\mathfrak{J}_2 \end{pmatrix} = \sum_{k=1}^{5} v_k\mathfrak{J}^{-1}\hat{\mathfrak{F}}_k, \tag{190}$$

we obtain

$$\hat{\mathfrak{G}} = \pm\mathfrak{G}[\mathfrak{M}] \tag{191}$$

or more explicitly

$$\begin{pmatrix} \hat{w}_4\mathfrak{J}_2 & \hat{\mathfrak{W}}\mathfrak{J}_2 \\ \mathfrak{J}_2\hat{\mathfrak{W}} & \hat{w}_5\mathfrak{J}_2 \end{pmatrix} = \pm\,\mathfrak{M}'\begin{pmatrix} w_4\mathfrak{J}_2 & \mathfrak{W}\mathfrak{J}_2 \\ \mathfrak{J}_2\mathfrak{W} & w_5\mathfrak{J}_2 \end{pmatrix}\mathfrak{M}. \tag{192}$$

On the other hand, if $\mathfrak{M}$ is any real matrix satisfying $\mathfrak{J}[\mathfrak{M}] = \pm\mathfrak{J}$, then the matrix $\hat{\mathfrak{G}}$ in (191) is again a skew-symmetric solution of (182) and consequently of the form (190), and (192) defines a real linear substitution $\hat{\mathfrak{w}} = \mathfrak{B}\mathfrak{w}$ with $\mathfrak{S}[\mathfrak{B}] = \mathfrak{S}$ and $|\mathfrak{B}| = 1$, by (188). We may replace $\mathfrak{M}$ by $-\mathfrak{M}$ and obtain the same matrix $\mathfrak{B}$, but the pair $\mathfrak{M}, -\mathfrak{M}$ is uniquely determined by $\mathfrak{B}$.

Imposing on $\mathfrak{B}$ the condition $\mathfrak{J}[\mathfrak{M}] = +\mathfrak{J}$, we obtain a subgroup of index 2 in the group of all $\mathfrak{B}$; and this subgroup is isomorphic to the inhomogeneous symplectic group, by (192). If we write

$$\mathfrak{M}' = \begin{pmatrix} \mathfrak{A} & \mathfrak{B} \\ \mathfrak{C} & \mathfrak{D} \end{pmatrix}$$

and calculate the single terms in (192), we find exactly the formulae (176), (177) and (178).

59. Let K be again a totally real algebraic number field of degree h. Let $\mathfrak{T}[\mathfrak{v}]$ be a quinary quadratic form with coefficients in K such that all conjugates except $\mathfrak{T}[\mathfrak{v}]$ itself are definite. We assume, moreover, that $\mathfrak{T}[\mathfrak{v}]$ has the signature 2, 3, i. e., that $\mathfrak{T}[\mathfrak{v}]$ can be transformed into $-v_1^2 + v_2^2 - v_3^2 + v_4^2 - v_5^2$ by a real linear substitution. Then there exists also a real matrix $\mathfrak{N}$, such that $\mathfrak{T}[\mathfrak{N}] = \mathfrak{S}$.

Consider now the units of $\mathfrak{T}$ in K, i. e., the integral matrices $\mathfrak{U}$ in K

satisfying $\mathfrak{T}[\mathfrak{U}] = \mathfrak{T}$. Since $|\mathfrak{U}| = \pm 1$ and $|-\mathfrak{U}| = -|\mathfrak{U}|$, we restrict ourselves to the case $|\mathfrak{U}| = +1$. Then the matrix $\mathfrak{R}^{-1}\mathfrak{U}\mathfrak{R} = \mathfrak{V}$ satisfies $\mathfrak{S}[\mathfrak{V}] = \mathfrak{S}$, $|\mathfrak{V}| = 1$, and the corresponding substitution (170) transforms H either into itself or into $\bar{H}$. We consider only the matrices $\mathfrak{U}$ with the first property; they form a subgroup of index 2 or 1 in the whole group of units with $|\mathfrak{U}| = 1$; we denote it by $\Lambda(\mathfrak{T})$, and by $\Delta(\mathfrak{T})$ the isomorphic subgroup of Ω.

Let us first prove that $\Delta(\mathfrak{T})$ is discontinuous. By Section **19**, it is sufficient to prove that $\Delta(\mathfrak{T})$ does not contain an infinite number of bounded elements. Otherwise $\Lambda(\mathfrak{T})$ would also contain infinitely many bounded matrices $\mathfrak{U}$, by (176), (177), (178) and $\mathfrak{U} = \mathfrak{N}\mathfrak{W}\mathfrak{N}^{-1}$. Since $\mathfrak{T}[\mathfrak{U}] = \mathfrak{T}$ and all conjugates of $\mathfrak{J}[\mathfrak{v}]$, except $\mathfrak{J}[\mathfrak{v}]$ itself, are definite, all conjugates of these matrices are bounded. Moreover their elements are integers, and this is a contradiction.

We apply now the following results from the theory of units of quadratic forms. Let Q be the space of real matrices $\mathfrak{Q}^{(32)}$ satisfying

$$\mathfrak{S}[\mathfrak{P}] > 0, \qquad \mathfrak{P}^{(52)} = \begin{pmatrix} \mathfrak{Q} \\ \mathfrak{E} \end{pmatrix}.$$

If

$$\mathfrak{V} = \begin{pmatrix} \mathfrak{V}_1 & \mathfrak{V}_2 \\ \mathfrak{V}_3 & \mathfrak{V}_4 \end{pmatrix}$$

is a real solution of $\mathfrak{S}[\mathfrak{V}] = \mathfrak{S}$, with $\mathfrak{V}_1 = \mathfrak{V}_1^{(3)}$, then the mapping

$$\mathfrak{Q} \to (\mathfrak{V}_1\mathfrak{Q} + \mathfrak{V}_2)(\mathfrak{V}_3\mathfrak{Q} + \mathfrak{V}_4)^{-1}$$

transforms Q into itself. We restrict $\mathfrak{V}$ to the matrices $\mathfrak{N}^{-1}\mathfrak{U}\mathfrak{N}$, where $\mathfrak{U}$ runs over the units of $\mathfrak{T}$ with $|\mathfrak{U}| = 1$. Then there exists in Q a fundamental domain Q_0, with respect to $\Lambda(\mathfrak{T})$, bounded by a finite number of algebraic surfaces and having only a finite number of neighbors. Moreover the integral

$$v(\mathfrak{T}) = \int_{Q_0} |\mathfrak{S}[\mathfrak{P}]|^{-5/2} d\mathfrak{Q} \tag{193}$$

has a finite value. As a matter of fact, a proof of these statements has been published only in the case of the field of rational numbers, instead of a totally real algebraic field K of arbitrary degree h; but the generalization of this proof offers no difficulties.

60. In order to derive the corresponding results for the group $\Delta(\mathfrak{T})$ and the space H, we have only to map Q onto H. If

(194) $$\mathfrak{Z} = \begin{pmatrix} z_1 & z_3 \\ z_3 & z_2 \end{pmatrix} \subset H, \qquad z_4 = |\mathfrak{Z}|, \qquad z_5 = 1,$$

we define

$$\mathfrak{z} = \mathfrak{x} + i\mathfrak{y}, \qquad \mathfrak{P} = (\mathfrak{x}\mathfrak{y}) \begin{pmatrix} x_4 & y_4 \\ 1 & 0 \end{pmatrix}^{-1} = \frac{1}{2iy_4^2} (\mathfrak{z}\bar{\mathfrak{z}}) \begin{pmatrix} 1 & -\bar{z}_4 \\ -1 & z_4 \end{pmatrix};$$

then the elements of $\mathfrak{Q} = (q_{kl})$ are

$$q_{k1} = \frac{y_k}{y_4}, \qquad q_{k2} = \frac{x_k y_4 - y_k x_4}{y_4} \qquad (k = 1, 2, 3)$$

and

$$\mathfrak{S}[\mathfrak{P}] = \frac{1}{2y_4^2} \mathfrak{S}\{\mathfrak{z}\} \mathfrak{E} \begin{bmatrix} 0 & y_4 \\ 1 & -x_4 \end{bmatrix} > 0;$$

hence $\mathfrak{Q}$ is a point of Q. On the other hand, if $\mathfrak{Q}$ is an arbitrary point of Q, we determine z_4 from the quadratic equation

(195) $$\mathfrak{S}[\mathfrak{P}] \begin{bmatrix} z_4 \\ 1 \end{bmatrix} = 0$$

and put

(196) $$\mathfrak{P} \begin{pmatrix} z_4 \\ 1 \end{pmatrix} = \mathfrak{z};$$

then

$$\mathfrak{S}[\mathfrak{z}] = 0, \qquad \mathfrak{S}\{\mathfrak{z}\} = \mathfrak{S}[\mathfrak{P}] \begin{Bmatrix} z_4 \\ 1 \end{Bmatrix} > 0.$$

If we replace z_4 by the other root $\bar{z}_4$ of (195), the column $\mathfrak{z}$ is replaced by $\bar{\mathfrak{z}}$; therefore we can choose the root z_4 of this quadratic equation, such that $y_1 > 0$, and then (194) is satisfied.

In this way, the fundamental domain Q_0 for $\Lambda(\mathfrak{T})$ in Q is mapped onto a fundamental domain F for $\Delta(\mathfrak{T})$ in H. It follows that F has only a finite number of neighbors $F\mathfrak{M}$, for the elements $\mathfrak{M}$ of $\Delta(\mathfrak{T})$, and that any compact domain in H is covered by a finite number of images $F\mathfrak{M}$; moreover F is bounded by a finite number of algebraic surfaces. Putting $\mathfrak{Z} = \mathfrak{X} + i\mathfrak{Y}$ and introducing the new variables $\mathfrak{X}$, $\mathfrak{Y}$ into (193), by (195) and (196), we obtain

$$v(\mathfrak{T}) = 4 \int_F |\mathfrak{Y}|^{-3} d\mathfrak{X} d\mathfrak{Y} = 4 \int_F d\mathfrak{X} d\mathfrak{Y}^{-1}.$$

This proves that F has a finite symplectic volume.

The latter result is not trivial in the case $h = 1$, since then H is not compact relative to $\Delta(\mathfrak{T})$; this may be derived from the theorem of A. Meyer, that an indefinite quinary quadratic form with rational coefficients is a zero form. On the other hand, $\mathfrak{T}[\mathfrak{v}]$ is not a zero form in K in the case $h > 1$,

since we assumed that the conjugates of $\mathfrak{T}[\mathfrak{v}]$ are definite except $\mathfrak{T}[\mathfrak{v}]$ itself; it can be proved as a simple consequence that then F is compact.

This is a sketch of the proof of Theorem 14; the details may be completed according to the scheme of Chapter VII.

Consider now the invariant subgroup $\Delta_\kappa(\mathfrak{T})$ of $\Delta(\mathfrak{T})$, defined by the condition $\mathfrak{U} \equiv \mathfrak{E} \pmod{\kappa}$ for the units $\mathfrak{U}$ of $\mathfrak{T}$, where κ denotes, as in Theorem 10, a certain power of an arbitrary prime ideal. It follows from Section **39**, that $\Delta_\kappa(\mathfrak{T})$ has no fixed points in H. The subgroup $\Delta_\kappa(\mathfrak{T})$ is of finite index j in $\Delta(\mathfrak{T})$, and the union of j images $F\mathfrak{M}$ of F, for suitably chosen elements $\mathfrak{M}$ of $\Delta(\mathfrak{T})$, constitutes a fundamental domain F_κ of $\Delta_\kappa(\mathfrak{T})$. Since the domain F_κ is compact in the case $h > 1$, it gives another example of a closed manifold with the symplectic metric.

It is known that the volume $v(\mathfrak{T})$ appears in the formula for the measure of the genus of $\mathfrak{T}$. In this way an analogue of Theorem 12 might be found.

61. We proved in Section **27**, that $\Delta(\mathfrak{G}, \mathfrak{H})$ is commensurable with a group $\Delta(r, s)$ of symplectic quaternion matrices $\pm \mathfrak{M}$. We shall now derive a corresponding result for the group $\Delta(\mathfrak{T})$.

A matrix of rank r is called *primitive*, if its elements are algebraic integers and if the minors of degree r are relative prime. Let $\mathfrak{W}_0^{(24)}$ be a primitive matrix of rank 2. Then the matrix $\mathfrak{F}[\mathfrak{W}_0] = \mathfrak{G}^{(4)} = (g_{kl})$ is skew-symmetric and has the rank 2; its elements are the minors of $\mathfrak{W}_0$. It follows by an application of Laplace's theorem, that also $\mathfrak{G}$ is primitive.

LEMMA 25. *Let $\mathfrak{G}^{(4)}$ be a primitive skew-symmetric matrix of rank 2, with elements from an algebraic number field K_0. Then there exists in K_0 a primitive matrix $\mathfrak{W}_0$, such that $\mathfrak{F}_0[\mathfrak{W}_0] = \mathfrak{G}$.*

Since $\mathfrak{G}$ is a primitive matrix of rank 2 and degree 4, with elements from K_0, we can determine in K_0 a unimodular matrix $\mathfrak{U}_0$, such that the first two columns of $\mathfrak{G}\mathfrak{U}_0$ are zero. Moreover the matrix $\mathfrak{G}[\mathfrak{U}_0]$ is skew-symmetric and again primitive; hence

$$\mathfrak{G}[\mathfrak{U}_0] = \epsilon \begin{pmatrix} 0 & 0 \\ 0 & \mathfrak{F}_2 \end{pmatrix},$$

where ϵ is an algebraic unit in K_0. Obviously we may choose $\mathfrak{U}_0$, such that $\epsilon = 1$. Denoting by $\mathfrak{W}_0$ the matrix of the two last rows of $\mathfrak{U}_0^{-1}$, we obtain the statement of the lemma.

For any skew-symmetric $\mathfrak{G}^{(4)} = (g_{kl})$,

$$|\mathfrak{G}| = (g_{12}g_{34} - g_{13}g_{24} + g_{14}g_{23})^2$$

and consequently

$$g_{12}g_{34} - g_{13}g_{24} + g_{14}g_{23} = 0, \tag{197}$$

if the rank of $\mathfrak{G}$ is < 4. It follows from (197) that the 36 minors of $\mathfrak{G}$, of degree 2, have the values $\pm g_{kl}g_{pq}$ $(1 \leq k < l \leq 4;\ 1 \leq p < q \leq 4)$. This proves that the skew-symmetric matrix $\mathfrak{G}^{(4)}$ is a primitive matrix of rank 2, if and only if (197) holds and the six numbers g_{kl} $(1 \leq k < l \leq 4)$ are relative prime.

Let $\mathfrak{W}_0 = (\mathfrak{C}_1^{(2)}\mathfrak{D}_1^{(2)})$ be the matrix of Lemma 25; then $\mathfrak{C}_1\mathfrak{D}'_1 - \mathfrak{D}_1\mathfrak{C}'_1 = (g_{13} + g_{24})\mathfrak{F}_2$, and consequently the relationship $\mathfrak{C}_1\mathfrak{D}'_1 = \mathfrak{D}_1\mathfrak{C}'_1$ holds, if and only if the condition

$$g_{13} + g_{24} = 0 \tag{198}$$

is satisfied. Since $\mathfrak{W}_0\mathfrak{U}_0 = (0\mathfrak{E})$, the equation $\mathfrak{A}_0\mathfrak{D}'_1 - \mathfrak{B}_0\mathfrak{C}'_1 = \mathfrak{E}$ has an integral solution $\mathfrak{A}_0, \mathfrak{B}_0$ in K_0. Then the matrices $\mathfrak{A}_1 = \mathfrak{A}_0 - \mathfrak{B}_0\mathfrak{A}'_0\mathfrak{C}_1$ and $\mathfrak{B}_1 = \mathfrak{B}_0 - \mathfrak{B}_0\mathfrak{A}'_0\mathfrak{D}_1$ satisfy

$$\mathfrak{A}_1\mathfrak{B}'_1 - \mathfrak{B}_1\mathfrak{A}'_1 = \mathfrak{B}_0\mathfrak{A}'_0(\mathfrak{C}_1\mathfrak{D}'_1 - \mathfrak{D}_1\mathfrak{C}'_1)\mathfrak{A}_0\mathfrak{B}'_0,$$
$$\mathfrak{A}_1\mathfrak{D}'_1 - \mathfrak{B}_1\mathfrak{C}'_1 = \mathfrak{E} - \mathfrak{B}_0\mathfrak{A}'_0(\mathfrak{C}_1\mathfrak{D}'_1 - \mathfrak{D}_1\mathfrak{C}'_1),$$

and the matrix

$$\mathfrak{M}_1 = \begin{pmatrix} \mathfrak{A}_1 & \mathfrak{B}_1 \\ \mathfrak{C}_1 & \mathfrak{D}_1 \end{pmatrix}, \tag{199}$$

with integral elements in K_0 and the second matrix row $\mathfrak{W}_0$, is symplectic, if (198) is fulfilled.

62. We consider again the substitution $\hat{\mathfrak{w}} = \mathfrak{V}\mathfrak{w}$ with $\mathfrak{S}[\hat{\mathfrak{w}}] = \mathfrak{S}[\mathfrak{w}]$ and $|\mathfrak{V}| = 1$. We assume now that the elements v_{kl} of $\mathfrak{V}$ are integers in a real algebraic field K_0 and that $\mathfrak{V} \equiv \mathfrak{E} \pmod 2$. We define

$$g_{14} = v_{51},\ g_{23} = -v_{52},\ g_{24} = \tfrac{1}{2}v_{53},\ g_{13} = -\tfrac{1}{2}v_{53},\ g_{12} = v_{54},\ g_{34} = v_{55}.$$

Then (198) is satisfied, and also (197), as a consequence of $\mathfrak{S}^{-1}[\mathfrak{V}'] = \mathfrak{S}^{-1}$. Since $|\mathfrak{V}| = 1$ and $v_{53} \equiv 0 \pmod 2$, the 6 numbers g_{kl} $(1 \leq k < l \leq 4)$ are relative prime. By the results of the last section, the matrix $\mathfrak{M}_1$ of (199) is symplectic with integral elements in K_0 and $\mathfrak{F}_2[\mathfrak{C}_1\mathfrak{D}_1] = \mathfrak{G} = (g_{kl})$, whence

$$\hat{w}_5 = \sum_{l=1}^{5} v_{5l}w_l = \sigma(\mathfrak{C}_1\mathfrak{W}\mathfrak{F}_2\mathfrak{D}'_1\mathfrak{F}_2^{-1}) + |\mathfrak{C}_1|\,w_4 + |\mathfrak{D}_1|\,w_5. \tag{200}$$

Let $\mathfrak{V}_1$ be the matrix of the linear transformation (176), (177), (178) with $\mathfrak{M}_1$ instead of $\mathfrak{M}$. By (178) and (200), the matrices $\mathfrak{V}$ and $\mathfrak{V}_1$ have the same fifth row, hence $(0\,0\,0\,0\,1)$ is the fifth row of $\mathfrak{W}\mathfrak{W}_1^{-1} = \mathfrak{W}_2$. Putting

$$\mathfrak{M}\mathfrak{M}_1^{-1} = \mathfrak{M}_2 = \begin{pmatrix} \mathfrak{A}_2 & \mathfrak{B}_2 \\ \mathfrak{C}_2 & \mathfrak{D}_2 \end{pmatrix},$$

we infer from (168) that

$$|\mathfrak{D}_2| = 1, \quad |\mathfrak{C}_2| = 0 \quad \text{and} \quad \sigma(\mathfrak{C}_2\mathfrak{W}\mathfrak{Z}_2\mathfrak{D}'_2\mathfrak{Z}_2^{-1}) = \sigma(\mathfrak{W}\mathfrak{D}_2^{-1}\mathfrak{C}_2) = 0,$$

whence $\mathfrak{C}_2 = 0$ and $\mathfrak{A}_2\mathfrak{D}'_2 = \mathfrak{E}$. The corresponding linear transformation (176) takes the simpler form

$$\hat{\mathfrak{W}} = \mathfrak{W}[\mathfrak{A}'_2] + w_5\mathfrak{B}_2\mathfrak{A}'_2. \tag{201}$$

Obviously $\mathfrak{B}_1$ is unimodular in K_0, hence the same is true for the matrix $\mathfrak{B}_2$, and the coefficients in (201) are integers. Let

$$\mathfrak{A}_2 = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

then a^2, b^2, c^2, d^2 are necessarily integers, hence also a, b, c, d. Moreover $\mathfrak{B}_2\mathfrak{A}'_2$ is integral and $|\mathfrak{A}_2| = 1$; hence $\mathfrak{B}_2$ is integral. This proves that $\mathfrak{M} = \mathfrak{M}_2\mathfrak{M}_1$ has integral elements. Since $|\mathfrak{M}| = 1$, these elements are relative prime.

On the other hand, by (176), (177) and (178), the elements of $\mathfrak{M}$ satisfy a system of algebraic equations with coefficients in K_0, and the only solutions of these equations are $\mathfrak{M}$ and $-\mathfrak{M}$. It follows that $\sqrt{t}\mathfrak{M} = \mathfrak{F}$, where $\mathfrak{F}$ is a matrix in K_0 and t is a number $\neq 0$ in K_0.

Since $\sqrt{t}$ is the greatest common divisor of the elements of $\mathfrak{F}$, the principal ideal (t) is the square of an ideal $\mathfrak{a}$ in K_0. Let $A_1 = E, A_2, \cdots, A_g$ denote the ambiguous classes of ideals in K_0, i. e., the classes A satisfying $A^2 = E$, where E is the principal class. We choose an integral ideal $\mathfrak{a}_k$ in A_k $(k = 2, \cdots, g)$ and take $\mathfrak{a}_1 = (1)$; then $\mathfrak{a}_k{}^2 = (a_k)$ is a principal ideal and $a_1 = 1$. Let $u_1, \cdots, u_s$ be a complete system of fundamental units in K_0 and denote by f_l $(l = 1, \cdots, 2^{s+1})$ the 2^{s+1} products $(-1)^{e_0}u_1{}^{e_1} \cdots u_s{}^{e_s}$ $(e_k = 0, 1;\ k = 0, \cdots, s)$, in particular $f_1 = 1$. The products $a_k f_l$ are all different; we denote them by $t_1, \cdots, t_m$ with $m = 2^{s+1}g$ and $t_1 = a_1 f_1 = 1$. Obviously $t = t_q v^2$, where t_q is one of the numbers $t_1, \cdots, t_m$ and v is a number in K_0.

None of the numbers $t_2, \cdots, t_m$ is a square in K_0. We choose now an integral ideal ω_0 in K_0, such that none of those numbers is a quadratic residue modulo ω_0. If $\mathfrak{B} \equiv \mathfrak{E} \pmod{2\omega_0}$, then $|\mathfrak{D}| \equiv 1$, $|\mathfrak{A}| \equiv 1$, $\mathfrak{C} \equiv 0$, $\mathfrak{B} \equiv 0$, by (176), (177) and (178). Moreover $\mathfrak{A}\mathfrak{D}' \equiv \mathfrak{E}$, and the coefficients of $\mathfrak{W}[\mathfrak{A}'] - \mathfrak{W}$ are $\equiv 0$. Putting

$$\mathfrak{A} = \begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix}, \qquad \mathfrak{D} = \begin{pmatrix} d_1 & d_2 \\ d_3 & d_4 \end{pmatrix},$$

we obtain

$$(202)\quad a_1^2 \equiv 1, \quad a_1 \equiv a_4 \equiv d_1 \equiv d_4, \quad a_2 \equiv a_3 \equiv d_2 \equiv d_3 \equiv 0 \pmod{2\omega_0}.$$

But $\sqrt{t_q}\, a_1 = x$ is an integer in K_0; hence $x^2 \equiv t_q \pmod{\omega_0}$, $t_q = t_1 = 1$.

We have proved that $\mathfrak{M}$ is an integral matrix in K_0, if $\mathfrak{B}$ is an integral matrix in K_0 satisfying $\mathfrak{S}[\mathfrak{B}] = \mathfrak{S}$, $|\mathfrak{B}| = 1$ and $\mathfrak{B} \equiv \mathfrak{E} \pmod{2\omega_0}$.

63. Let $\mathfrak{T}[\mathfrak{v}]$ be the quinary quadratic form of Section **59**. There exists in K a matrix $\mathfrak{L}_0$, such that

$$\mathfrak{T}[\mathfrak{L}_0\mathfrak{v}] = m(-pv_1^2 + v_2^2 - qv_3^2 + rsv_4^2 - rv_5^2)$$

with integral positive m, p, q, r, s. Putting

$$(pqs)^{-1} = m_2, \quad pm_2 = m_1, \quad qm_2 = m_3, \quad rsm_2 = m_4, \quad rm_2 = m_5,$$

we have

$$(203)\quad \begin{gathered} \prod_{k=1}^{5} m_k = r^2 m_2^4, \\ -m^{-1}m_2\mathfrak{T}[\mathfrak{L}_0\mathfrak{v}] = \sum_{k=1}^{5} (-1)^{k-1} m_k v_k^2. \end{gathered}$$

Let $\mathfrak{L}_1$ be the matrix of the linear substitution

$$(204)\quad \begin{gathered} w_1 = -v_1\sqrt{m_1} - v_2\sqrt{m_2}, \quad w_2 = v_1\sqrt{m_1} - v_2\sqrt{m_2}, \quad w_3 = -v_3\sqrt{m_3}, \\ w_4 = -v_4\sqrt{m_4} - v_5\sqrt{m_5}, \quad w_5 = v_4\sqrt{m_4} - v_5\sqrt{m_5} \end{gathered}$$

and let $\mathfrak{L}_0\mathfrak{L}_1^{-1} = \mathfrak{N}_1$; then $m^{-1}m_2\mathfrak{T}[\mathfrak{N}_1] = \mathfrak{S}$.

By (185) and (186), the matrices

$$\mathfrak{Q}_k = \sqrt{m_k}\, \mathfrak{J}_k \qquad (k = 1, \cdots, 5)$$

satisfy the conditions

$$(205)\quad \begin{gathered} \mathfrak{Q}_k\mathfrak{Q}_l = -\mathfrak{Q}_l\mathfrak{Q}_k \ (1 \le k < l \le 5), \quad \mathfrak{Q}_k^2 = (-1)^{k-1} m_k \mathfrak{E} \ (k = 1, \cdots, 5), \\ \prod_{k=1}^{5} \mathfrak{Q}_k = r m_2^2 \mathfrak{E}. \end{gathered}$$

Using the abbreviations

$$\mathfrak{J}_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \mathfrak{J}_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \mathfrak{J}_3 = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix},$$
$$\sqrt{p} = \rho_1, \quad \sqrt{s} = \rho_2, \quad \sqrt{q} = \rho_3, \quad \sqrt{r} = \rho_4,$$

we introduce the 16 linearly independent matrices

$$\mathfrak{P}_1 = \begin{pmatrix} \mathfrak{E} & 0 \\ 0 & \mathfrak{E} \end{pmatrix}, \qquad \mathfrak{P}_2 = m_2^{-1}\mathfrak{Q}_1\mathfrak{Q}_2 = \rho_1 \begin{pmatrix} -\mathfrak{F}_3 & 0 \\ 0 & \mathfrak{F}_3 \end{pmatrix},$$

$$\mathfrak{P}_3 = m_2^{-1}s\mathfrak{Q}_1\mathfrak{Q}_2\mathfrak{Q}_3 = \rho_2 \begin{pmatrix} -\mathfrak{E} & 0 \\ 0 & \mathfrak{E} \end{pmatrix}, \qquad \mathfrak{P}_4 = ps\mathfrak{Q}_3 = \rho_1\rho_2 \begin{pmatrix} \mathfrak{F}_3 & 0 \\ 0 & \mathfrak{F}_3 \end{pmatrix},$$

$$\mathfrak{P}_5 = m_2^{-1}\mathfrak{Q}_2\mathfrak{Q}_3 = \rho_3 \begin{pmatrix} -\mathfrak{F}_1 & 0 \\ 0 & \mathfrak{F}_1 \end{pmatrix}, \qquad \mathfrak{P}_6 = m_2^{-1}\mathfrak{Q}_1\mathfrak{Q}_3 = \rho_1\rho_3 \begin{pmatrix} \mathfrak{F}_2 & 0 \\ 0 & \mathfrak{F}_2 \end{pmatrix},$$

$$\mathfrak{P}_7 = qs\mathfrak{Q}_1 = \rho_2\rho_3 \begin{pmatrix} \mathfrak{F}_1 & 0 \\ 0 & \mathfrak{F}_1 \end{pmatrix}, \qquad \mathfrak{P}_8 = m_2^{-1}\mathfrak{Q}_2 = \rho_1\rho_2\rho_3 \begin{pmatrix} -\mathfrak{F}_2 & 0 \\ 0 & \mathfrak{F}_2 \end{pmatrix},$$

$$\mathfrak{P}_9 = m_2^{-1}\mathfrak{Q}_1\mathfrak{Q}_3\mathfrak{Q}_4 = \rho_4 \begin{pmatrix} 0 & -\mathfrak{E} \\ -\mathfrak{E} & 0 \end{pmatrix}, \qquad \mathfrak{P}_{10} = m_2^{-1}p\mathfrak{Q}_2\mathfrak{Q}_3\mathfrak{Q}_4 = \rho_1\rho_4 \begin{pmatrix} 0 & -\mathfrak{F}_3 \\ \mathfrak{F}_3 & 0 \end{pmatrix},$$

$$\mathfrak{P}_{11} = m_2^{-1}\mathfrak{Q}_2\mathfrak{Q}_4 = \rho_2\rho_4 \begin{pmatrix} 0 & \mathfrak{E} \\ -\mathfrak{E} & 0 \end{pmatrix}, \qquad \mathfrak{P}_{12} = m_2^{-1}\mathfrak{Q}_1\mathfrak{Q}_4 = \rho_1\rho_2\rho_4 \begin{pmatrix} 0 & \mathfrak{F}_3 \\ \mathfrak{F}_3 & 0 \end{pmatrix},$$

$$\mathfrak{P}_{13} = m_2^{-1}q\mathfrak{Q}_1\mathfrak{Q}_2\mathfrak{Q}_4 = \rho_3\rho_4 \begin{pmatrix} 0 & \mathfrak{F}_1 \\ -\mathfrak{F}_1 & 0 \end{pmatrix}, \qquad \mathfrak{P}_{14} = pq\mathfrak{Q}_4 = \rho_1\rho_3\rho_4 \begin{pmatrix} 0 & \mathfrak{F}_2 \\ \mathfrak{F}_2 & 0 \end{pmatrix},$$

$$\mathfrak{P}_{15} = m_2^{-1}\mathfrak{Q}_3\mathfrak{Q}_4 = \rho_2\rho_3\rho_4 \begin{pmatrix} 0 & -\mathfrak{F}_1 \\ -\mathfrak{F}_1 & 0 \end{pmatrix}, \qquad \mathfrak{P}_{16} = m_2^{-2}\mathfrak{Q}_1\mathfrak{Q}_2\mathfrak{Q}_3\mathfrak{Q}_4 = \rho_1\rho_2\rho_3\rho_4 \begin{pmatrix} 0 & -\mathfrak{F}_2 \\ \mathfrak{F}_2 & 0 \end{pmatrix}.$$

Then any real matrix $\mathfrak{M}^{(4)}$ can be expressed in the form

$$\mathfrak{M} = \sum_{k=1}^{16} x_k \mathfrak{P}_k \tag{206}$$

with uniquely determined real scalar factors x_k.

We denote the real algebraic number field $K(\sqrt{p}, \sqrt{q}, \sqrt{r}, \sqrt{s})$ by K_0. Let ω_0 be the ideal of the preceding section and choose in K an ideal ω having the corresponding properties with respect to K instead of K_0. We put $4pqrs\omega\omega_0 = \mu$.

On account of (203) and (204), the elements of the matrix $\mathfrak{N}_1$ lie in K_0. Let f be an integer $\neq 0$, such that $f\mathfrak{N}_1$ and $f\mathfrak{N}_1^{-1}$ are integral, and choose in K an ideal ν which is divisible by $f^2\mu$. We consider now the elements $\mathfrak{U}$ of the congruence subgroup $\Lambda_\nu(\mathfrak{T})$ of $\Lambda(\mathfrak{T})$ defined by condition $\mathfrak{U} \equiv \mathfrak{E} \pmod{\nu}$. Then $\mathfrak{V} = \mathfrak{N}_1^{-1}\mathfrak{U}\mathfrak{N}_1$ is an integral matrix in K_0 satisfying $\mathfrak{S}[\mathfrak{V}] = \mathfrak{S}$, $|\mathfrak{V}| = 1$ and $\mathfrak{V} \equiv \mathfrak{E} \pmod{\mu}$. The corresponding symplectic matrices $\pm\mathfrak{M}$ form a subgroup $\Delta_\nu(\mathfrak{T})$ of $\Delta(\mathfrak{T})$.

By Section **58**, the pair $\mathfrak{M}, -\mathfrak{M}$ is uniquely determined by the conditions

$$\hat{\mathfrak{Q}}_k = \mathfrak{M}^{-1}\mathfrak{Q}_k\mathfrak{M} \ (k = 1, \cdots, 5), \qquad \mathfrak{F}[\mathfrak{M}] = \mathfrak{F} \tag{207}$$

with

$$\hat{\mathfrak{Q}}_k = \sum_{l=1}^{5} r_{lk}\mathfrak{Q}_l \ (k = 1, \cdots, 5), \qquad (r_{kl}) = \mathfrak{L}_0^{-1}\mathfrak{U}\mathfrak{L}_0. \tag{208}$$

On account of the result of Section **62**, the coefficients x_k in (206) are numbers of K_0. We apply any isomorphism of K_0 which leaves all numbers of K

invariant and denote by $\mathfrak{P}^*_k$, $\mathfrak{Q}^*_k$, $\hat{\mathfrak{Q}}^*_k$, x^*_k the images of $\mathfrak{P}_k$, $\mathfrak{Q}_k$, $\hat{\mathfrak{Q}}_k$, x_k. Then (205) and (208) hold good with $\mathfrak{Q}^*_k$, $\hat{\mathfrak{Q}}^*_k$ instead of $\mathfrak{Q}_k$, $\hat{\mathfrak{Q}}_k$, and the same is true for the relationship

$$\mathfrak{Q}'_k = \mathfrak{J}^{-1}\mathfrak{Q}_k\mathfrak{J} = (\mathfrak{Q}_2\mathfrak{Q}_4)^{-1}\mathfrak{Q}_k(\mathfrak{Q}_2\mathfrak{Q}_4) \qquad (k = 1, \cdots, 5).$$

It follows that both matrices

$$\mathfrak{M}^* = \sum_{k=1}^{16} x^*_k\mathfrak{P}^*_k, \qquad \mathfrak{M}_0 = \sum_{k=1}^{16} x_k\mathfrak{P}^*_k$$

are solutions of (207), with $\mathfrak{Q}^*_k$, $\hat{\mathfrak{Q}}^*_k$ instead of $\mathfrak{Q}_k$, $\hat{\mathfrak{Q}}_k$, and consequently

$$\mathfrak{M}^* = \pm\mathfrak{M}_0, \qquad x^*_k = \pm x_k \qquad (k = 1, \cdots, 16).$$

Putting

$$\mathfrak{M} = \begin{pmatrix} \mathfrak{A} & \mathfrak{B} \\ \mathfrak{C} & \mathfrak{D} \end{pmatrix},$$

we have

$$\mathfrak{A} = \begin{pmatrix} x_1 + x_2\rho_1 - x_3\rho_2 - x_4\rho_1\rho_2 & (-x_5 + x_6\rho_1 + x_7\rho_2 - x_8\rho_1\rho_2)\rho_3 \\ (-x_5 - x_6\rho_1 + x_7\rho_2 + x_8\rho_1\rho_2)\rho_3 & x_1 - x_2\rho_1 - x_3\rho_2 + x_4\rho_1\rho_2 \end{pmatrix},$$

$$\mathfrak{D} = \begin{pmatrix} x_1 - x_2\rho_1 + x_3\rho_2 - x_4\rho_1\rho_2 & (x_5 + x_6\rho_1 + x_7\rho_2 + x_8\rho_1\rho_2)\rho_3 \\ (x_5 - x_6\rho_1 + x_7\rho_2 - x_8\rho_1\rho_2)\rho_3 & x_1 + x_2\rho_1 + x_3\rho_2 + x_4\rho_1\rho_2 \end{pmatrix};$$

if we replace x_k by $(-1)^k x_{k+8}\rho_4$ $(k = 1, \cdots, 8)$, we obtain the expressions for $\mathfrak{B}$, $\mathfrak{C}$. By (176), (177) and (178), $\mathfrak{B} \equiv 0 \pmod{\mu}$ and $\mathfrak{C} \equiv 0 \pmod{\mu}$, hence $x_9, \cdots, x_{16}$ are integers. Moreover (202) holds now for the module μ instead of $2\omega_0$; consequently $x_1, \cdots, x_8$ are also integers and $x_1^2 \equiv 1 \pmod{\omega}$. Since $|\mathfrak{M}| = 1$, the numbers $x_1, \cdots, x_{16}$ are relative prime. We apply the argument of Section **62** and conclude that the coefficients x_k are numbers of the field K.

Let $\Delta(p, q, r, s)$ denote the group of all symplectic matrices $\pm\mathfrak{M}$ with integral $x_1, \cdots, x_{16}$ in K. We have proved that $\Delta_\nu(\mathfrak{T})$ is a subgroup of $\Delta(p, q, r, s)$. On the other hand, if $\mathfrak{M}$ is an element of $\Delta(p, q, r, s)$ satisfying $x_1 \equiv 1$, $x_k \equiv 0 \pmod{f^2}$ for $k = 2, \cdots, 16$, then $\mathfrak{V} \equiv \mathfrak{E} \pmod{f^2}$, by (176), (177) and (178), and $\mathfrak{U} = \mathfrak{N}_1\mathfrak{V}\mathfrak{N}_1^{-1}$ is an element of the group $\Lambda(\mathfrak{T})$. Consequently the groups $\Delta(\mathfrak{T})$ and $\Delta(p, q, r, s)$ are commensurable.

INSTITUTE FOR ADVANCED STUDY,
PRINCETON, N. J.

THE THEORY OF SIMPLE RINGS.*

By Emil Artin and George Whaples.

Introduction. In this paper we discuss the theory of simple associative rings; for such rings we prove structure theorems which are generalizations of the well known theorems concerning commutators, automorphisms, etc. of simple algebras.[1] We use a new method—namely, the study of the transformations which the functions

$$\sigma(x) = \sum_{\nu=1}^{\eta} a_\nu x b_\nu$$

produce on the elements of the ring. The reader will be able, if he wishes, to extract from our discussion proofs of the structure theorems for algebras which are shorter than those usually given. For his convenience, we have occasionally inserted known proofs, in order to make the theory complete.

1. Vector spaces over rings. If R is a ring, we call a set V of elements an R-vector space if its elements form an abelian group under addition and if R is an operator domain on this abelian group, satisfying the usual axioms—namely the distributive law and the condition that multiplication in the ring is the same as combination of operators. Usually R operates on the left side, and in that case the product of operators corresponds to the function-theoretic product $(\sigma\tau(x) = \sigma(\tau(x)))$; but occasionally we write the operators on the other side, and then this is not true. In the following, we shall usually assume that no non-zero vector is annihilated by every element of R. The only spaces we shall consider will be R-subspaces.

A space is said to satisfy the *minimal condition* if any set of R-subspaces contains a smallest one. This is equivalent to the descending chain condition; that is, to the condition that every descending chain of subspaces must be finite. The minimal condition on spaces is occasionally easy to verify by the following:

Theorem 1. *If V is the sum (not necessarily direct sum) of the subspaces V_1 and V_2, both of which satisfy the minimal condition, then V satisfies the minimal condition.*

* Received January 26, 1942.

[1] For these theorems, and for complete references, see M. Deuring, *Algebren*, Ergebnisse der Mathematik, vol. 4, no. 1, Berlin, 1935, or A. A. Albert, *Structure of Algebras*, New York, 1939.

Proof. Let W be any subspace of V, W_2 the intersection of W and V_2, and W_1 the set of all elements v_1 which appear in any possible equation $w = v_1 + v_2$. (These elements will be called the V_1-components of W.) Then both W_1 and W_2 are subspaces of V. Let W' be a subspace of W. We contend:

LEMMA. *If W' contains W_2, and if all of W_1 occurs among the V_1-components of W', then $W' = W$.*

Proof. Let $w = v_1 + v_2$ be any element of W. Then v_1 lies in W_1, so that there is an element w' with the same V_1-component: $w' = v_1 + v'_2$. Since $w - w' = v_2 - v'_2$ and belongs to W, it belongs to W_2. Therefore $w - w'$ belongs to W' and so w itself belongs to W'.

Now suppose we have a set of subspaces W of $V_1 + V_2$. Consider first all those of our set for which the corresponding subset W_2 of V_2 is minimal, and, among these, one for which W_1 is minimal. There obviously cannot be a smaller W in our set.

A space V is said to be irreducible if $RV \neq 0$ and if V contains no proper subspaces other than 0.

LEMMA. *If the R-space V is the sum of r irreducible subspaces $V_1, V_2, \cdots V_r$, and if $x_1, x_2, \cdots x_s$ are any $s > r$ vectors of V, then one of them can be written as a linear combination of the others, with coefficients in R.*

Proof. The lemma is evident for $r = 0$, because that means that V is the null-space. Assume that it is true for $r - 1$ and use induction. Let

$$\begin{aligned} x_1 &= x_{11} + x_{12} + \cdots + x_{1r} \\ x_2 &= x_{21} + x_{22} + \cdots + x_{2r} \\ &\cdots\cdots\cdots \\ x_s &= x_{s1} + x_{s2} + \cdots + x_{sr} \end{aligned}$$

be a splitting of our vectors according to the covering $V = V_1 + V_2 + \cdots + V_r$. If the first term of each of these sums is 0, then $x_1, x_2, \cdots x_s$ are contained in the subspace $V_2 + V_3 + \cdots + V_r$, and the lemma follows from our assumption. Otherwise we can assume $x_{11} \neq 0$. Then Rx_{11} is a non-zero subspace of V_1 and so it follows that $Rx_{11} = V_1$. We can, therefore, find elements $a_2, a_3, \cdots a_s$ in R, such that $a_i x_{11} = x_{i1}$. The new elements

$$y_2 = x_2 - a_2 x_1; \qquad y_3 = x_3 - a_3 x_1; \cdots; \qquad y_s = x_s - a_s x_1$$

are then obviously contained in the subspace $V_2 + V_3 + \cdots + V_s$. Since their number $s - 1 > r - 1$, one of them, say y_s, can be expressed in terms of the others. From this expression, we easily get an expression for x_s in terms of $x_1, x_2, \cdots x_{s-1}$.

A sum $V = V_1 + V_2 + \cdots + V_n$ is said to be direct whenever every element of V is uniquely expressible as a sum of elements of the V_i, or, what is the same thing, whenever $v_1 + v_2 + \cdots + v_n = 0$ implies that each $v_i = 0$. In case our sum is not direct and V_n is irreducible, there is obviously only this alternative: the sum of the spaces V_n and $(V_1 + V_2 + \cdots + V_{n-1})$ is direct, or else V_n has a vector in common with the sum of the others; then V_n, since it is irreducible, is included in this sum, and $V = V_1 + V_2 + \cdots + V_{n-1}$. Should all the V_i be irreducible, we should therefore get a direct sum by removing as many terms as possible from our expression $V = V_1 + V_2 + \cdots + V_n$; the sum which remains will be direct.

The following assumption gives a useful class of special R-spaces: Let every vector of V be contained in a finite sum of irreducible subspaces. We shall call such spaces *regular*.

If V is a regular space and contains a direct sum $V_1 + V_2 + \cdots + V_r$ of irreducible subspaces, we have the following alternative: either V is equal to that sum, or there is a vector of V not contained in it. Since the space is regular, there must, then, be an irreducible space V_{r+1} not contained in our previous sum; then $V_1 + V_2 + \cdots + V_{r+1}$ is direct. Should this process never come to an end by leading to the full space, we should obtain an infinite sequence $V_1, V_2, V_3, \cdots$ of irreducible subspaces with a direct sum. The following descending chain

$$\begin{aligned} W_1 &= V_1 + V_2 + V_3 + \cdots \\ W_2 &= \quad\; V_2 + V_3 + \cdots \\ W_3 &= \quad\quad\; V_3 + \cdots \\ &\;\;\cdot \quad\;\; \cdot \quad\;\; \cdot \end{aligned}$$

shows that in such a case the minimal condition cannot be fulfilled. A regular space with minimal condition is consequently the direct sum of a finite number of irreducible spaces. Conversely, the direct sum of a finite number of irreducible spaces is, by Theorem 1, a space with minimal condition. The number r of components in such a direct splitting is obviously an invariant, for if V is the direct sum of $W_1, W_2, \cdots W_s$, it follows from our lemma that s cannot be larger than r; if it were, we could select vectors different from 0 in $W_1, W_2, \cdots W_s$ such that one of them is expressible in terms of the others, contradicting the directness of the sum.

If V and W are two R-spaces, a homomorphism of V to W is a mapping of each vector of V onto some vector of W, which is a homomorphism for the group operation and is connected with R by the assumption that if $\sigma(v) = w$, then $\sigma(rv) = rw$, for all r in R. The set of all images of V need not fill out the whole space W, but it is always a subspace. Should this subspace be all

of W, then we may say that W is the homomorphic image of V. On the other hand, many vectors of V might be mapped onto 0 of W. The set of all these forms a subspace of V; should this subspace be 0, and should W be the homomorphic image of V, then the mapping is really a one to one correspondence and the two spaces are said to be isomorphic.

We now readily see that in case V and W are irreducible, we have only this alternative: either all of V is mapped onto 0, or the mapping is an isomorphism. Let $V = V_1 + V_2 + \cdots V_r$ be a direct splitting of V into irreducible subspaces, and let W be a subspace of V. The following procedure will give us a homomorphic mapping of W onto V_1: split any element w according to the direct sum into components, and map w onto its V_1-component v_1. If we, furthermore, assume W to be irreducible, we see that if V_1 and W are not isomorphic, our homomorphic mapping must map W onto 0, and consequently the V_1-component of any element of W is 0. So in this case W is already contained in the direct sum of those among $V_1, V_2, \cdots V_r$ which are isomorphic to W. This suggests a grouping of the components V_i in such a way that isomorphic V_i are united. Should $V_1 + V_2 + \cdots + V_s$ be the space generated by such a group, then we see that it contains all irreducible subspaces W of V which are isomorphic to V_1, and no others. We can, therefore, describe this space in a way not related to any special splitting, by saying that it is the sum of all subspaces isomorphic to V_1. That its multiplicity s is invariant follows from our previous discussion.

Every ring R can always be considered as an R-space, by defining the operator product of the operator a and the vector x (where of course both of them are in R) as the ordinary product ax. A subspace, under this concept of R as left R-space, is then a left ideal. R can also be considered as right R-space; the subspaces of R are then right ideals. Since it is very useful to study these spaces, it is clear that we shall have to make some finiteness assumption about R: we usually assume, therefore, that R satisfies the minimal condition for left ideals, or, what is the same thing, that R is a left R-space with minimal condition.

2. Simple rings. A simple ring is a ring R containing no two sided ideal other than itself and 0, and such that $R^2 \neq 0$. A vector space over a simple ring is either annihilated by the whole ring, or it is not annihilated by any element other than 0; for the annihilators of an R-space V form a two sided ideal, since $aV = 0$ implies both $RaV = 0$ and $aRV = a(RV) \subset aV = 0$. This shows that a non-zero left ideal I of a simple ring has no left annihilator other than 0; for since the set of annihilators of I is a two sided ideal, the only other possibility is that I is annihilated by the whole ring R. But then

R is annihilated on the right by I, and then, for the same reason, the set of right annihilators of R would be R itself, contradicting $R^2 \neq 0$. The same reasoning applies to right ideals.

Now assume that R contains minimal left ideals (ideals containing no subideal other than 0). Let V be an irreducible space and I a minimal left ideal. Since $IV \neq 0$, we can find a vector x_0 such that $Ix_0 = V$. If we map the element λ of I onto λx_0, we obviously get a homomorphism; since both spaces are irreducible, it is an isomorphism. Hence all irreducible spaces are isomorphic, and the isomorphism with any minimal left ideal can be produced by the process described. Let us especially identify V with I. x_0 is then an element of I such that $\lambda \rightarrow \lambda x_0$ is a one to one correspondence of I onto itself. Since x_0 must appear among the images, there is an element e of I such that $ex_0 = x_0$. Hence $ex_0 = e^2x_0 \neq 0$, and the fact that it is a one to one correspondence shows that $e^2 = e \neq 0$. If I is a minimal left ideal, then Ia is a homomorphic image of I, and hence is either 0 or isomorphic to I. Let us now consider the sum of all minimal left ideals;[2] from the preceding sentence, it follows that this sum is a two-sided ideal and consequently is the whole ring R.

We can now prove a theorem concerning the existence of a unit element.

THEOREM 2. *If R is simple and contains minimal ideals, it contains a unit element if and only if it satisfies the minimal condition for left ideals.*

Proof. R is the sum of its minimal left ideals. Suppose it contains a unit element. Then this element 1 can be expressed as the sum $\lambda_1 + \lambda_2 + \cdots \lambda_r$ of a finite number of components, each of which is contained in a minimal ideal. The sum $R\lambda_1 + R\lambda_2 + \cdots + R\lambda_r$ therefore contains $R \cdot 1 = R$, and each $R\lambda_i$ is a minimal ideal. So R is covered by a finite number of irreducible spaces and hence satisfies the minimal condition.

To show the other part of the theorem, we first prove a more general statement:

LEMMA. *If R is a simple ring with left minimal condition and* I *is any left ideal, then there is an idempotent e in* I *such that* $I = Re$.

Proof. Since every ideal contains a minimal ideal, it also contains some idempotent. Among all the idempotents of I select one, e_1, such that the set of elements of I annihilated by e_1 from the right is as small as possible. (This set always forms a left ideal $\mathfrak{m}$.) We contend that $\mathfrak{m} = 0$. If $\mathfrak{m}$ were

[2] A sum of an infinite number of sets is defined to be the set of all finite sums of elements of the components.

not 0, it would contain some idempotent e_2 different from 0; then $e_2e_1 = 0$. Put $e = e_1 - e_1e_2 + e_2$; then $ee_1 = e_1$, $ee_2 = e_2$ and so $ee = ee_1 - ee_1e_2 + ee_2 = e$. Now $xe = 0$ implies $xee_1 = xe_1 = 0$, so that e can never annihilate more than e_1. But e really annihilates less than e_1, because $e_2e = e_2$, whereas $e_2e_1 = 0$.

Thus I contains an idempotent which annihilates no non-zero element from the right; let e be such an element. If x is any element of I, $x - xe$ is also in I and is annihilated from the right by e; hence $x = xe$. We see that e acts as a right unit for all elements of I. If we apply the lemma to the case $\mathrm{I} = R$, we get the existence of a right unit e in R. The solutions of the equation $ex = 0$ would form a right ideal r. Now $\mathrm{r} \cdot \mathrm{r} = (\mathrm{r}e)\mathrm{r} = \mathrm{r}(e\mathrm{r}) = 0$; since a right ideal is annihilated from the right only by 0, $\mathrm{r} = 0$. But $x - ex$ is annihilated by e from the left; hence $x = ex$.

COROLLARY. *A simple ring with minimal condition for left ideals satisfies the minimal condition for right ideals. In fact, if Re is a minimal left ideal, eR is a minimal right ideal.*

This corollary will follow from Theorem 2 as soon as we have proved the following

LEMMA. *In any simple ring R, an ideal of the form Re (e idempotent) is minimal if and only if eRe is a quasi-field.*

Proof. Let R be a simple ring and Re a minimal ideal. For any non-zero element a of eRe, we know that $Ra \neq 0$, since R contains e and $ea = a$. So $Ra = Re$; this implies $eRa = eRe$ and, since $ea = a$, $eRea = eRe$. Hence the ring eRe has a unit element, and a left inverse of any non-zero element.

If, conversely, a is any non-zero element of Re, the set eRa is a left ideal in eRe, since $eRe \cdot eRa \subset eRa$. If, therefore, eRe is a quasi-field, eRa can be only 0 or eRe. In the first case the left ideal Ra would be annihilated by e on the left; this can happen in a simple ring only if Ra is 0. In the other case, we conclude that $ReRa = ReRe$. Since $ReR = R$, this gives us $Ra = Re$; Re is minimal.

Thus if R is simple with minimal condition for left ideals, R contains a minimal right ideal and has a unit element; hence, by Theorem 2, (applied to right ideals instead of left ideals) R has the minimal condition for right ideals.

THEOREM 3. *If $\mathrm{I} = Re$ is a minimal ideal, then its R-homomorphisms are exactly the set of transformations obtained by multiplying I on the right by elements of eRe; hence, if these homomorphisms are written as right operators, they form a ring isomorphic to eRe.*

Proof. $\mathfrak{I} \cdot eRe = Re \cdot eRe = Re$, so that multiplication of $\mathfrak{I}$ by any element of eRe gives a homomorphic mapping of $\mathfrak{I}$ on itself. If, conversely, a is the image of e under any homomorphic mapping, the image of xe will be xa. Setting x equal to e, we get $ea = a$; since a is in $\mathfrak{I}$, we have also $a = ae$. Hence $a = eae \in eRe$.

THEOREM 4. *To every simple ring R, with minimal condition, we can find a quasi-field k and a right space of finite dimension over that quasi-field, such that R is isomorphic to the set of all k-linear transformations of that space, if we write them on the left side. The space will be any minimal left ideal and the field the set of all R-homorphisms of that left ideal.*

Conversely, the set of all k-linear transformations of any right space of finite dimension over a quasi-field k is a simple ring R with minimal condition; the space is irreducible, and therefore isomorphic to any minimal left ideal of R; and the field is isomorphic to the set of R-homomorphisms of the space.

Thus the structure of k and the dimension of the space are uniquely determined by R.[3]

Proof. We split R into a direct sum of minimal right ideals: $R = \mathfrak{r}_1 + \mathfrak{r}_2 + \cdots + \mathfrak{r}_n$. If

$$1 = e_1 + e_2 + \cdots + e_n \tag{1}$$

is the corresponding splitting of the unit element 1, we multiply (1) by e_i from the right and compare the result with the obvious splitting of e_i into components: $e_i = 0 + \cdots + 0 + e_i + 0 + \cdots + 0$. This shows that e_i is idempotent and that $e_i e_k = 0$ if $i \neq k$. Multiplying (1) from the right by R shows that none of the e_i can be 0, or else we could split R into fewer right ideals.

We now take $\mathfrak{I} = Re_1$ as our left ideal, and let $k = e_1 R e_1$. e_i cannot annihilate $\mathfrak{I}$ from the left, because a left ideal has no left annihilator. Select $\alpha_i \neq 0$ out of $e_i\mathfrak{I}$. Then $\alpha_i e_1 R e_1 = \alpha_i R e_1$; but $\alpha_i R$ is part of $e_i R$ and not 0, so that $\alpha_i k = e_i R e_1 = e_i\mathfrak{I}$. Hence $\alpha_1 k + \alpha_2 + \cdots + \alpha_n k = e_1\mathfrak{I} + e_2\mathfrak{I} + \cdots + e_n\mathfrak{I} \supset \mathfrak{I}$ and is therefore equal to $\mathfrak{I}$.

If we multiply $\alpha_1, \alpha_2, \cdots \alpha_n$ on the left side by xe_i, α_k is taken into 0 if $k \neq i$, and α_i is taken into $x\alpha_i$. Since $R\alpha_i = \mathfrak{I}$, we see that α_i can be transformed into an arbitrary element of $\mathfrak{I}$ while the other α_k are taken into 0. This shows that the α_i are linearly independent over k and that we can find an element $x \in R$ which takes the n vectors α_i into arbitrarily given images in $\mathfrak{I}$.

---

[3] This theorem was first proved (for simple algebras) by J. H. M. Wedderburn: "On hypercomplex numbers," *Proceedings of the London Mathematical Society*, vol. 6 (1908), pp. 77-118.

To prove the second half of the theorem, let V be a right k-space of dimension n, and R the ring of all k-homomorphisms of V. Let $\mathfrak{a}$ be a two sided ideal of R and $a \neq 0$ an element of $\mathfrak{a}$. Select any n independent vectors $A_1, A_2, \cdots A_n$ of V. a cannot take all these vectors into 0; assume that it maps A_1 onto $B \neq 0$. Let σ be a transformation mapping A_i onto A_1 and all the other A_j onto 0; then $a\sigma$ will map A_i onto B and all the other basis vectors onto 0. Let τ be a transformation mapping B onto C; $\tau a\sigma$ will map A_i onto an arbitrary vector C and all the other A_j onto 0. But $\tau a\sigma$ also belongs to $\mathfrak{a}$; it follows that $\mathfrak{a} = R$.

That V is irreducible over R is clear, since every vector different from 0 can be mapped onto any other vector. It remains to be shown that if we consider V as R-space only, then the field k consists of all R-homomorphisms. Let A be a fixed vector different from 0 and let B be its image under a given R-homomorphism of V. If B were linearly independent of A, with respect to k, we could find an element σ of R such that $\sigma(A) = A$ and $\sigma(B) = 0$. Our homomorphism mapped A onto B, so it should map $\sigma(A)$ onto $\sigma(B)$, i. e. A onto 0, contrary to assumption. Consequently, A is taken into $A\kappa$. If C is any other vector, we can find a σ such that $\sigma(A) = C$; then $\sigma(A\kappa) = C\kappa$. Our homomorphism takes A into $A\kappa$, and therefore takes $\sigma(A)$ into $\sigma(A\kappa)$; hence it takes any vector C into $C\kappa$, with the same κ.

Any space over a simple ring with minimal condition is regular. Indeed, any vector x is contained in Rx and therefore in $\mathfrak{l}_1 x + \mathfrak{l}_2 x + \cdots + \mathfrak{l}_n x$, where the $\mathfrak{l}_i$ are minimal ideals. But every $\mathfrak{l}_i x$ is either 0 or irreducible.

If two spaces with minimal condition over a simple ring R with minimal condition split into the same number of irreducible components, they are isomorphic. The isomorphism consists, indeed, in mapping each irreducible component of V onto one of the irreducible components of V', and then extending that mapping to a mapping of V on V' by addition. If we have a slightly more general situation, namely, two rings R and R', that are isomorphic under the mapping $r \leftrightarrow r'$, and an R-space V and an R'-space V', which both split into the same number of irreducible components, we can do something similar: we first consider V' as an R-space, by defining the multiplication of V' by an element r of R to have the same effect as the multiplication by the image r' of r in R'. Hence there is a function σ mapping V onto V' in such a manner that for any vector v, and any element r of R, we have $\sigma(rv) = r\sigma(v)$. The multiplication of v by r is the original one; since $\sigma(v)$ is in V' the multiplication by r means really the multiplication by r'. We consequently have

THEOREM 5. *If R and R' are two simple, isomorphic rings which contain*

minimal ideals, and V and V' are spaces over R and R', respectively, splitting into the same number of irreducible components, we can find a mapping function $\sigma(v)$ of V onto V', which is a one to one correspondence and satisfies the identity $\sigma(rv) = r'\sigma(v)$ (where r' is the image of r under our isomorphism).

COROLLARY. *Under the same conditions as before, except for the equality of the number of irreducible components of the two spaces with minimal condition, it is possible to map one of the spaces, say V, onto a part of the other, by a one to one correspondence $\sigma(v)$ such that the same equation holds.*

Proof. Should V' have more components than V, we can find a subspace of V' having the same number.

3. Direct products. We give a definition of direct product of two spaces over a quasi-field k.

If k is a quasi-field and A a right, B a left k-space (such that the unit of k acts as unit on A and B), we consider the set of all finite formal sums:

$$\sum_{\nu} A_\nu B_\nu \tag{1}$$

with $A_\nu \in A$, $B_\nu \in B$. We allow at first no simplification whatever besides the commutative law of addition, and are neither allowed to cancel two terms nor to drop a term with a factor 0. Two such sums are added in the obvious way. These sums do not form an additive group, because we have neither a zero element nor a negative. The associative and commutative laws are obvious.

Such a sum shall be called a nullform if and only if there is a substitution

$$B_i = \sum_{\mu} \beta_{i\mu} B'_\mu \tag{2}$$

for the B_i by some other vectors B'_i of $B(\beta_{i\mu} \in k)$ such that we get 0 as coefficient of each B'_i if we formally substitute (2) in (1) and collect all the coefficients of B'_i—that is, if for all i:

$$\sum_{\nu} A_\nu \beta_{\nu i} = 0. \tag{3}$$

The set of all nullforms shall be denoted by N_B. If (1) is a nullform such that (2) and (3) hold, we select any linearly independent vectors A'_i, such that every A_i can be expressed by them:

$$A_i = \sum A'_\mu \alpha_{\mu i}. \tag{4}$$

Substituting this in (3), we have $\sum_{\mu,\nu} A'_\mu \alpha_{\mu\nu} \beta_{\nu i} = 0$. Since the A'_i are linearly independent, we get:

$$\sum \alpha_{k\nu} \beta_{\nu i} = 0 \tag{5}$$

for all k and i. Multiplying (5) from the right by B'_i and taking the sum over i, we see that because of (2):

$$\sum \alpha_{kv} B_v = 0. \tag{6}$$

But (6) means that (4) substituted in (1) leads to a coefficient 0 of A'_i if we collect terms. This shows the symmetry of our definition of nullforms; namely, $N_B \subset N_A$ (where N_A stands for the nullforms defined by A-substitution.) Consequently, $N_A = N_B$, so that we denote this set merely by N. Furthermore, we have shown that the question whether or not (1) is a null form can be decided by expressing the A_i in terms of *any* set of linearly independent A'_i. Since our definition of N is now known to be symmetric with regard to the A's and B's, we have:

LEMMA. *The decision whether* (1) *is a nullform can be reached by expressing the* B_i *in terms of any set of linearly independent vectors* B'_i: (1) *is a nullform if and only if* (3) *holds.*

It is fairly obvious how to prove that the sum of two nullforms is a nullform, and that $\Sigma(-A_v)B_v$ is in N if (1) is in N. But we need also:

LEMMA. *If* $\Sigma A_v B_v + \Sigma \bar{A}_v \bar{B}_v \in N$ *and* $\Sigma \bar{A}_v \bar{B}_v \in N$, *then* $\Sigma A_v B_v \in N$.

Express both the B_v and the $\bar{B}_v$ in terms of a set of linearly independent B'_v. These expressions are uniquely determined: $B_i = \Sigma \beta_{iv} B'_v$, $\bar{B}_i = \Sigma \bar{\beta}_{iv} B'_v$. Since $\Sigma \bar{A}_v \bar{B}_v \in N$, we have $\Sigma \bar{A}_v \bar{\beta}_{vi} = 0$. From this and the first of our assumptions, we deduce $\Sigma A_v B_v \in N$.

Finally, we remark that $AB + (-A)B$ is always in N. Indeed, both the B-vectors can be expressed by one alone, namely B. The coefficient of this B after collecting is $A - A = 0$.

If X stands for a sum (1), we agree to let $-X$ stand for $\Sigma(-A_v)B_v$.

We now introduce a congruence:

$$X \equiv Y \pmod{N}$$

means that $X + (-Y) \in N$. Reflexiveness and symmetry are fairly obvious. We concentrate on transitivity. Let

$$X + (-Y) \in N; \qquad Y + (-Z) \in N.$$

Then

$$X + (-Z) + Y + (-Y) \in N.$$

Since $Y + (-Y) \in N$, we deduce $X + (-Z) \in N$ from the second of our lemmas.

If $X \equiv Y \pmod{N}$ and $Z \equiv W \pmod{N}$, we have $X + Z \equiv Y + W \pmod{N}$. In the sense of congruence, our forms are now a group with every element of N as zero element and $-X$ as negative of X. The rules

$$\begin{aligned} A(B + B') &\equiv AB + AB' \pmod{N} \\ (A + A')B &\equiv AB + A'B \pmod{N} \\ A(\alpha B) &\equiv (A\alpha)B \pmod{N} \end{aligned}$$

show that in the sense of congruence any formal change is allowed with the vectors A and B.

We now substitute $=$ for $\equiv$ and call this abelian group $A \times_k B$, the direct product of the spaces A and B relative to k.

If A and B are rings as well as k-spaces, and k is a field operating on both sides of A and B, we define the product of the elements $\sum_\nu a_\nu b_\nu$ and $\sum_\mu a'_\mu b'_\mu$ to be $\sum_{\nu,\mu} (a_\nu a'_\mu)(b_\nu b'_\mu)$. To show that this defines a ring, it suffices to show that this multiplication is well defined relative to equality in $A \times_k B$. To show this, we need only prove that if $\Sigma a_\nu b_\nu$ is a nullform, then $\Sigma (a a_\nu)(b b_\nu)$ is a nullform. But if the substitution $b_i = \Sigma \beta_{i\mu} b'_\mu$ makes the first sum obviously 0, the substitution $(b b_i) = \Sigma \beta_{i\mu} (b b'_\mu)$ makes the second obviously 0. So $A \times_k B$ is a ring. Clearly $A \times_k B$ is isomorphic to $B \times_k A$; we shall allow the order of the factors of the terms of (1) to be changed and say that $A \times_k B = B \times_k A$: the direct product is commutative. It is also easy to verify that $A \times_k (B \times_k C) = (A \times_k B) \times_k C$: direct product is associative. If A and B are of finite degree over k, then the degree of $A \times_k B$, over k, is the product of the degrees of A and of B, since if $a_1 \cdots a_n$ and $b_1 \cdots b_m$ are linearly independent bases for A and for B, the elements $a_i b_j$ are a linearly independent basis for $A \times_k B$.

If B has a unit element 1_B, the elements $a 1_B$ of $A \times_k B$ form a subring isomorphic to A. Indeed, $a 1_B = 0$ if and only if $a = 0$, since 1_B is linearly independent. The properties for addition and multiplication are obvious. We identify this subring with A itself. Should A have a unit element, we can in a similar way identify B with the subring of elements $1_A b$. Since $(a 1_B)(1_A b) = ab$, we see furthermore that in that sense $A \times_k B$ is really the product of A and B.

If R is a simple ring, we call any function $l(x)$ of the form $\Sigma a_\nu x a'_\nu$, where x is a variable in R, an analytic linear function. If l and l' are two analytic linear functions, their iteration $l(l'(x))$ is obviously again analytic. The main theorem about analytic linear functions in a simple ring is the following:

THEOREM 6. *If R is a simple ring with unit element, and centrum k, and if $a_1, a_2, \cdots a_r$ are linearly independent with respect to k, we can find an analytic linear function mapping these elements on arbitrarily chosen elements $b_1, b_2, \cdots b_r$.*

Proof. It obviously suffices to prove the theorem in case one of the b_i is 1 and the others 0. If $r = 1$, the possible values $l(a_1)$ form the two sided ideal Ra_1R, which is R because R is simple: this proves our theorem for $r = 1$. Let it be true for $r - 1$ and assume that we have found, for each $i = 1 \cdots r - 1$, an analytic linear function $l_i(x)$ that maps a_i into 1 and all the others (except possibly a_r) into 0. We try to find an $l(x)$ which maps a_r into 1 and all the other a_i into 0. This is easily done as soon as any $l_i(a_r)$ is an element outside the centrum k; indeed, if c is not commutative with $l_i(a_r)$, the linear function $\lambda(x) = l_i(x)c - cl_i(x)$ will map $a_1, a_2, \cdots a_{r-1}$ into 0 and a_r into a non-zero element b. If $\mu(x)$ is a linear function mapping b into 1, we can take $\mu(\lambda(x))$ as our desired function $l(x)$.

We therefore assume that $l_i(a_r) = \kappa_i \in k$ for $i = 1 \cdots r - 1$. We now put $\lambda(x) = a_1l_1(x) + a_2l_2(x) + \cdots + a_{r-1}l_{r-1}(x) - x$. Obviously $\lambda(a_i) = 0$ for $i = 1, \cdots, r - 1$. $\lambda(a_r)$ cannot be 0 for else $0 = a_1\kappa_1 + a_2\kappa_2 \cdots + a_{r-1}\kappa_{r-1} - a_r$, contradicting the linear independence of the a_i with respect to k. If $\mu(x)$ maps $\lambda(a_r)$ into 1, the function $l(x) = \mu(\lambda(x))$ is the solution.

THEOREM 7. *If A is a simple ring with unit element and centrum k, and B is any ring with unit element, containing k in its centrum, then every two sided ideal of $A \times_k B$ is of the form $A \times_k \mathfrak{b}$, where $\mathfrak{b}$ is a two sided ideal of B.*

Proof. Let $\mathfrak{A}$ be any ideal of $A \times_k B$, and $\mathfrak{b}$ the intersection of $\mathfrak{A}$ with B. It is clear that $A \times_k \mathfrak{b}$ is part of $\mathfrak{A}$. Let, conversely, $\xi = \Sigma a_\nu b_\nu$ be an element of $\mathfrak{A}$. We can assume that all the a_ν are linearly independent with respect to k, since we could otherwise simplify the expression for ξ. If $l_i(x)$ is an analytic linear function with coefficients in A, mapping a_i onto 1 and all other a's onto 0, we find that $l_i(\xi) = b_i$. Since $l_i(\xi)$ belongs, with ξ, to the two sided ideal $\mathfrak{A}$, every b_i belongs to $\mathfrak{b}$, and hence $\mathfrak{A} = A \times_k \mathfrak{b}$.

COROLLARY. *If A and B are simple, with unit element, and the centrum k of A is contained in the centrum of B, then $A \times_k B$ is simple.*

With the help of the preceding theorem we can, in many cases, determine the intertwining relations of two subrings A and B of a bigger ring R.

THEOREM 8. *Let A and B be two subrings of any ring R, with the following properties:*

A and B both have the same unit element (which need not be unit element for R).

A is simple, and its centrum k is included in the centrum of B.

Every element of B is commutative with every element of A.

Then the subring $A \cdot B$ of R has the structure $A \times_k B$. This means that there are no intertwining relations between A and B other than the most trivial ones.

Proof. We construct two new rings $\bar{A}$ and $\bar{B}$, isomorphic to A and B, and having only the image $\bar{k}$ of k as common elements. We form the product $\bar{A} \times_{\bar{k}} \bar{B}$ and contend that the following mapping onto $A \cdot B$ is an isomorphism: let $\Sigma \bar{a}_\nu \bar{b}_\nu$ correspond to $\Sigma a_\nu b_\nu$. We have first to prove that the mapping is uniquely determined, but if $\Sigma \bar{a}_\nu \bar{b}_\nu = 0$, there is a transformation in the $\bar{b}_i$ that makes this obvious. The corresponding transformation in $\Sigma a_\nu b_\nu$ will show that this element is also 0, because the computation involves only the distributive and associative laws. That the mapping is a homomorphism under addition and multiplication is also obvious. It remains to be seen that it is a one to one correspondence. The set $\mathfrak{A}$ of all elements of $\bar{A} \times_{\bar{k}} \bar{B}$ that are mapped onto 0 forms a two sided ideal. According to our theorem, this ideal has the structure $\bar{A} \times_{\bar{k}} \bar{\mathfrak{b}}$, where $\bar{\mathfrak{b}}$ consists of those elements $\bar{b}$ of $\bar{B}$ which are mapped onto 0. Consequently $\bar{\mathfrak{b}} = 0$, and therefore $\mathfrak{A} = 0$.

This allows us to determine the structure of the set of all analytic linear functions of R, from another point of view. These functions form themselves a ring (with iteration used as multiplication). This ring, L, contains two subrings: First, the functions $l_a(x) = ax$. The set of all these functions is clearly isomorphic to R as soon as R contains a unit element, and we can, without misunderstanding, identify this subring with R itself. Second, the functions $l_a^*(x) = xa$. The set of all these functions is not isomorphic to R, because

$$l_a^*(l_b^*(x)) = xba = l_{ba}^*(x).$$

But the mapping of R onto the l_a^* is an anti-isomorphism, interchanging the order in a product. This new ring will be called R^*; its structure is of course known if the structure of R is known. If R is simple, with unit element, R^* has the same properties. l_a and l_b^* are commutative, since $l_a l_b^*(x)) = axb = l_b^*(l_a(x))$. If, finally, κ is in the centrum of R, $l_\kappa(x) = l_\kappa^*(x)$. This shows that the two rings satisfy all conditions of the preceding theorem, and therefore $R \times_k R^*$ is contained in L. Since, on the other hand, every element of L can be obtained from the special functions R and R^* by iteration and addition, we have $L = R \times_k R^*$. This shows that L is again simple.

As for the minimal condition in direct products, we prove:

THEOREM 9. *If A and B have unit elements included in the field k, the minimal condition in a ring $A \times_k B$ is satisfied if it is true for A and B is of finite degree over k.*

Proof. The minimal condition means that $A \times_k B$, considered as left space over $A \times_k B$, satisfies the minimal condition for subspaces. Now each $A \times_k B$-space is certainly an A-space. If, therefore, we prove the minimal condition for $A \times_k B$, considered as A-space, then our theorem follows. Let $w_1, w_2, \cdots w_n$ be a basis of B with respect to k. Then $A \times_k B = Aw_1 + Aw_2 + \cdots + Aw_n$. Each of the spaces Aw_i is isomorphic to A, so the minimal condition for $A \times_k B$ follows by use of Theorem 1.

Coming back to the linear functions $L = R \times_k R^*$ of a simple ring with centrum k, we can prove:

THEOREM 10. *The minimal condition for left ideals of L is satisfied if and only if R is of finite degree over k.*

It suffices to prove the second half. Let, therefore, R not be of finite degree over k, and $w_1, w_2, w_3, \cdots$ an infinite sequence of linearly independent elements (with respect to k).

Let $\mathfrak{I}_i$ stand for the set of all analytic linear functions $l_i(x)$ that map $w_1, w_2, \cdots w_i$ onto 0. Clearly $\mathfrak{I}_i \supset \mathfrak{I}_{i+1}$, and differs from it, because $\mathfrak{I}_i$ contains a linear function that maps $w_1, \cdots w_i$ onto 0 and w_{i+1} onto 1. Since these are left ideals, we have proved the theorem.

If R is of finite dimension over k, the structure of L is completely determined by Theorem 6:

THEOREM 11. *If R is simple, and of finite degree over its centrum k, then $L = R \times_k R^*$ consists of all linear transformations of the space R, with respect to k.*

Proof. Since every analytic linear transformation satisfies the conditions $l(x+y) = l(x) + l(y)$ and $l(\kappa x) = \kappa l(x)$, for every element κ of the centrum, L is part of the set of all linear transformations. If, on the other hand, $w_1, w_2, \cdots w_n$ are a linearly independent basis for B, with respect to k, we can find a function $l(x)$, mapping $w_1, w_2, \cdots w_n$ on arbitrarily selected images. But any linear transformation with respect to k is uniquely determined by the images of the w_i.

Another way to state the previous theorem is this:

If R is simple and of finite degree n over k, the ring $R \times_k R^*$ is isomorphic to the ring of all matrices of degree n over k.

4. Commutators of subrings. If R is a ring with minimal condition, and we want to study a subring of L that also satisfies the minimal condition, it seems natural to restrict one of the two factors of $L = R \times_k R^*$, by replacing it by a simple subring A of finite dimension over k. Such a ring A is called a simple algebra. We shall therefore study the subring $S = R \times_k A^*$. It again satisfies the minimal condition, and consists of the linear functions

$$l(x) = \Sigma r_\nu x a_\nu, \tag{1}$$

with $r_\nu \in R$ and $a_\nu \in A$. We might now consider R as a left S-space, in the sense that l is applied onto x, as left operator, by computing $l(x)$.

Since S contains the linear functions denoted by R, every S-space is also a left R-space; since R satisfies the minimal condition over R, the same is true when R is considered as space over S. Assume now that we have in R a second algebra A', isomorphic to A, under an isomorphism $a \leftrightarrow a'$ that leaves every element of k fixed. The set $S' = R \times_k A'^*$ is then isomorphic to S under an isomorphism that maps the function (1) onto $l'(x) = \Sigma r_\nu x a'_\nu$. We remark that the special function $l(x) = xa$ is mapped onto the function xa'. Let us now make use of the corollary to Theorem 5. R appears as space to the two isomorphic rings S and S', and we may therefore assume that a function $\sigma(x)$ can be found, which maps, by a one to one correspondence, R onto some part of itself, and satisfies, for every $l \in S$, the equation

$$\sigma(l(x)) = l'(\sigma(x)). \tag{2}$$

We first specialize to $l(x) = rx$ and get $\sigma(rx) = r\sigma(x)$. Putting $x = 1$, and writing c for $\sigma(1)$, we have $\sigma(r) = rc$. Since r can be any element of R, we have $\sigma(x) = xc$. This mapping is a one to one correspondence between R and the images, so that $xc = 0$ implies $x = 0$. c being no right divisor of 0, it has an inverse.[4] Specializing now in (2) to $l(x) = xa$, we get $xac = xca'$. Putting $x = 1$, we find:

$$a' = c^{-1}ac.$$

This gives us the following:

THEOREM 12. *Let R be a simple ring with minimal condition and A and A' two simple subalgebras of R containing the centrum k of R. Any isomorphism of A to A', that leaves every element of the centrum k fixed, can be extended to an inner automorphism of R.*

---

[4] The left ideals Rc^ν form a descending chain; hence, for some n, $Rc^{n+1} = Rc^n$. In the set on the right we have c^n and therefore, for some x, $xc^{n+1} = c^n$; since c is no divisor of zero, $xc = 1$.

COROLLARY. *If R is an algebra, every automorphism of R that leaves the centrum fixed is inner.*

Proof. Select $A = A' = R$.

In order to study the way in which a simple subring is embedded in a larger ring, we need to define the degree of a ring, not only over a subfield, but over a simple subring. To do this, we commence with the:

LEMMA. *Let S be a simple subring, with minimal condition, of any ring R, and let R and S have the same unit element. If $\mathfrak{I}_1$ and $\mathfrak{I}_2$ are two minimal ideals of S, then any isomorphism of $\mathfrak{I}_1$ onto $\mathfrak{I}_2$, considered as S-spaces, can be extended to an isomorphism of $R\mathfrak{I}_1$ to $R\mathfrak{I}_2$, considered as R-spaces.*

Proof. If e_1 and e_2 are idempotents of $\mathfrak{I}_1$ and $\mathfrak{I}_2$, let c_2 be the image of e_1 under the mapping of $\mathfrak{I}_1$ onto $\mathfrak{I}_2$. Then $x = xe_1$ of $\mathfrak{I}_1$ is mapped onto xc_2 of $\mathfrak{I}_2$. If we multiply, therefore, the set $Re_1 = RSe_1 = R\mathfrak{I}_1$ from the right by c_2, this mapping is an R-homomorphism and an extension of our mapping of $\mathfrak{I}_1$ onto $\mathfrak{I}_2$. The set of images is $Rc_2 = RSc_2 = R\mathfrak{I}_2$.

Let now c_1 be the image of e_2 under the inverse of our isomorphism, so that $x \in \mathfrak{I}_2$ is hereby mapped onto xc_1. Then $c_2c_1 = e_1$. If now $xe_1c_2 = 0$, we would have $xe_1c_2c_1 = xe_1 = 0$. Our mapping is therefore an isomorphism.

Now let $S = \mathfrak{I}_1 + \mathfrak{I}_2 + \cdots + \mathfrak{I}_r$ be a splitting in minimal left ideals and $1 = \Sigma e_\nu$ be the corresponding splitting in orthogonal idempotents. We easily deduce $R = Re_1 + Re_2 + \cdots + Re_r$. This is a direct sum, since $x_1e_1 + x_2e_2 + \cdots + x_re_r = 0$ shows after right multiplication by e_i that $x_ie_i = 0$. We now consider R as a left S-space. Each of the components Re_i is then also a left S-space, and they are isomorphic, since they are even isomorphic if considered as R-spaces. Therefore we can break each of them into an equal number of irreducible S-spaces. This will lead, after regrouping, to a splitting of R into a finite or infinite sum of left S-spaces, each of which is a direct sum of precisely r components, and is thus isomorphic to S itself. An S-subspace of R that is isomorphic to S has the following structure: let w be the image of 1 if we map S onto the subspace. x of S is then mapped onto xw of the subspace, and since the mapping is an isomorphism, $xw = 0$ with x in S implies $x = 0$. Doing this with each of our spaces, we see that there exist elements w_i, such that $R = \Sigma Sw_i$. Any finite number of w_i is linearly independent with respect to S, in the sense that a linear combination of them, with coefficients in S, is 0 if and only if all coefficients (not merely all terms) of the sum are 0. The number of irreducible S-components of the S-space R is r times as large as the number of w_i; therefore the number of w_i is invariant; we shall call it the left degree of R over S and shall denote it by $(R:S)$. Now

in precisely the same way as in case of fields, we can prove that $(R:S_1) = (R:S)(S:S_1)$, if S_1 is a simple subring of S.

If R is a ring and S a subset of R, we denote by R^S the set of all elements of R which are commutative with every element of S. If S is still a subset of R, and T another ring with unit element, we have $(R \times_k T)^S = R^S \times_k T$. For the right side is certainly contained in the left. Let $\Sigma r_\nu t_\nu$ be any element of the left side, with linearly independent t_ν. If s is any element of S, then $\Sigma(sr_\nu - r_\nu s)t_\nu = 0$. Since the t_ν are linearly independent, we see that every r_ν is in R^S. Similarly, if $S \subset R$ and $U \subset T$, and S and T are rings containing k, $(R \times_k T)^{S \times_k U} = R^S \times_k T^U$. R^R is of course the centrum of R.

Now let R be a ring and T be an algebra, both with centrum k. Then $(R \times_k T)^{R \times_k T} = k \times_k k = k$, so that k is also the centrum of $R \times_k T$. Assume now that R and T are simple rings with minimal condition, that each contains a simple algebra containing k, and that these two algebras are isomorphic under a correspondence leaving the elements of k invariant. We shall denote both these algebras, though different, by the same letter A, since it will be clear from the formulas which is meant. The simple ring with minimal condition $R \times_k T$ now contains the algebra A in two ways; consequently there is an inner automorphism of $R \times_k T$ that will map one onto the other. Computing $(R \times_k T)^A$ for both algebras A, we get the two sets $R^A \times_k T$ and $R \times_k T^A$. These are, therefore, isomorphic under the same inner automorphism. Computing further

$$(R \times_k T)^{(R^A \times_k T)} \text{ and } (R \times_k T)^{(R \times_k T^A)}$$

we get the two sets $R^{(R^A)}$ and $T^{(T^A)}$, which are therefore isomorphic under the same isomorphism.

Should we be able to prove that R^A is simple, we could compute the degree of $R \times_k T$ over $R^A \times_k T$; since an inner automorphism of $R \times_k T$ maps $R^A \times_k T$ onto $R \times_k T^A$, the degree of $R \times_k T$ over $R^A \times_k T$ should be the same as the degree of $R \times_k T$ over $R \times_k T^A$. The reader can easily deduce the equality $(R:R^A) = (T:T^A)$. The proof for the simplicity will be given soon.

If we only know that our algebra A is contained in R, we have to construct a T. Assume that the degree $(A:k)$ is n. Consider A as a k-space, and denote by M the set of all linear transformations of A over k; that is, the algebra of $n \times n$ matrices over k. The degree $(M:k)$ will be n^2. M contains a subalgebra isomorphic to A, namely the set of all linear transformations of A that are obtained by multiplying A from the left by an element of A. Calling this ring of linear transformations also A, we now ask for M^A, i. e. for the linear transformations σ that satisfy the condition $\sigma(ax) = a\sigma(x)$, for all a

and x in A. Putting $x = 1$, denoting $\sigma(1)$ by c, and writing x instead of a, we see: $\sigma(x) = xc$. It follows easily that the set of these linear transformations is isomorphic to A^*, so that we have $M^A = A^*$. (We did not assume that the centrum of A is k; the centrum might be larger, so that M would possibly not be equal to $A \times_k A^*$.)

Substituting M for T in our previous result, we can prove:

THEOREM 13. *If R is simple with minimal condition, and A is a simple subalgebra of R containing the centrum k of R, and if also $(A:k) = n$ and M is the set of all linear transformations of A over k, we have:*

1) *$R^A \times_k M$ and $R \times_k A^*$ are isomorphic under an inner automorphism of $R \times_k M$.*
2) *R^A is simple.*
3) *$R^{(R^A)} = A$.*
4) *$(R:R^A) = n$.*

Proof. The first statement is clear. Since A^* is simple, $R \times_k A^*$ is simple. This proves that $R^A \times_k M$, and consequently R^A, is simple. As for the third statement, $R^{(R^A)}$ is mapped by our inner automorphism onto $M^{(M^A)} = M^{A^*}$, which is the A contained in M. Since we know that our isomorphism maps precisely the A of R onto the A of M, we get our statement. The fourth statement follows from the fact that $(M:A^*) = n$. (For $(R:R^A) = (R \times_k M : R^A \times_k M) = (R \times_k M : R \times_k A^*) = (M:A^*)$). This follows in turn from $(M:k) = n^2$ and $(A^*:k) = n$.

There is one very general theorem concerning simple subalgebras of rings, that holds for any ring R whose centrum contains the centrum k of A and whose unit element is contained in k. The proof of this theorem depends on the following:

LEMMA. *Assume that an analytic linear function $l(x) = \Sigma a_\nu x a'_\nu$, with coefficients in A, is 0 for all $x \in A$. Then $l(x)$ is 0 for all $x \in R$.*

Proof. Since we can express the a'_ν by linearly independent elements of A and collect terms, we can from the beginning assume that the a'_ν are linearly independent. The necessary operation for collecting is the interchange of x with an element of the centrum k of A; this can be done, according to our assumption, for all $x \in R$.

We contend that now all the a_ν are 0. To show this, we consider $l(x)$ for values of A only. The linear functions of A form a ring isomorphic to $A \times_k A^*$, and under this isomorphism our linear function will correspond to

$\Sigma a_\nu a'^*_\nu$. Since $l(x) = 0$ for all $x \in A$, we have $\Sigma a_\nu a'^*_\nu = 0$, and the linear independence of the a' shows that all the a_ν are 0. But this means that $l(x) = 0$ for every x in R.

Our first aim is to compute R^A; we contend that this can be done in the following way:

THEOREM 14. *Let $l(x)$ be any analytic linear function with coefficients in A which is not identically zero and which maps every element of A onto an element of k. The values $l(x)$, for all x in R, will then range precisely over R^A.*

Proof. If a is any element of A, $l(x)a - al(x)$ is a linear function with coefficients in A that is 0 for every $x \in A$. According to our lemma, it is 0 for every x in R. Since a is arbitrary, this proves that $l(x) \in R^A$.

To prove that we get every element of R^A, take any element a such that $l(a) = \kappa \neq 0$. Since we can replace a by $\kappa^{-1}a$, we can assume that $l(a) = 1$. If z is any element of R^A, we get $l(za) = zl(a) = z$, because we can interchange z with every coefficient.

THEOREM 15. *Under the same assumptions as before, we have $R = R^A \times_k A$.*

Proof. Let $a_1, a_2, \cdots a_n$ be a k-basis for A. Let $l_i(x)$ be the analytic linear function with coefficients in A that takes a_i into 1 and all the other a's into 0. The linear function $l(x) = a_1 l_1(x) + a_2 l_2(x) + \cdots + a_n l_n(x) - x$ is 0 for $x = a_1, a_2, \cdots a_n$, and therefore for every x in A. This proves that for every x in R we have $x = a_1 l_1(x) + a_2 l_2(x) + \cdots + a_n l_n(x)$. Since the values of $l_i(x)$ range through R^A, we see that $R = A \cdot R^A$. That the dot can be replaced by the cross follows from Theorem 8.

COROLLARY. *If A is a simple algebra with unit element e and centrum k, contained in a ring R (for which we make no assumption), then $eR^k = A \times_k eR^A$.*

Proof. Consider first the ring $(eRe)^k$. Obviously the centrum of this ring contains k. e is unit element for this ring, and the previous lemma can be applied. But $((eRe)^k)^A = (eRe)^A = eR^A$.

THEOREM 16. *Let R be simple with chain condition, and let k be its centrum and T a subring containing k. Then if $R = w_1 T + w_2 T + \cdots w_n T$, R^T is an algebra A, and $(A : k) \leq n$.*

Furthermore, these two statements are equivalent:

1) *T is the commutator of a simple subalgebra of R.*
2) *T is simple, $(R:T)$ is finite, and all homomorphisms of R into itself, considering R as right T-space, are analytic.*

Proof. Let T be any subring of R such that $R = w_1T + w_2T + \cdots + w_nT$. Consider the set S of all homomorphisms of R, considered as right T-space. This ring S contains as subring the analytic transformations among them. Let $\sigma(x) = \Sigma r_\nu x r'_\nu$, and assume the r_ν linearly independent with respect to k. Since, for every $t \in T$, $\sigma(xt) - \sigma(x)t = 0$, we get the equations $tr'_\nu - r'_\nu t = 0$, so that all r'_ν belong to $A = R^T$. This subring is therefore to be denoted by $R \times_k A^*$. It contains as further subring the left multiplications of x by elements of R, and this ring of transformations we have always denoted by R. Since the degree $(A:k) = (R \times_k A^*:R)$, our inequality will be proved if we estimate the degree $(S:R)$. Every homomorphism $\sigma(x)$ of S will map $w_1, w_2, \cdots w_n$ onto certain elements $a_1, a_2, \cdots a_n$ of R. If these elements are known and $x = w_1t_1 + w_2t_2 + \cdots + w_nt_n$, then its image will be $a_1t_1 + a_2t_2 + \cdots + a_nt_n$. Let us denote $\sigma(x)$ by the n-dimensional vector $(a_1, a_2, \cdots a_n)$. We are interested in the behavior of S, considered as left R-space, where R means here the transformations $\rho(x) = rx$. Iterating our previous σ with that transformation would lead to $(ra_1, ra_2, \cdots ra_n)$. This shows that S is isomorphic to a certain part of the set of all n-dimensional vectors over R. The statement concerning the degrees can also be interpreted as a statement about the number of irreducible components: the degree has only to be multiplied by the number of irreducible components R has. This makes it clear that $(S:R) \leq n$. (From previous work, we know that this degree exists). Hence $(A:k) \leq n$. It is furthermore clear that $(S:R) = n$ if and only if S fills out the whole of our space of n-vectors. This means that there are linear transformations mapping the w_i onto arbitrary images. But then the w_i have to be linearly independent, because $w_1t_1 + w_2t_2 + \cdots + w_nt_n = 0$ would yield $t_i = 0$ if we applied the homomorphism mapping w_i onto 1 and the others onto 0. If, conversely, the w_i are linearly independent, and we map $w_1t_1 + w_2t_2 + \cdots + w_nt_n$ onto $a_1t_1 + a_2t_2 + \cdots + a_nt_n$, this clearly gives an element of S.

If A is a simple subalgebra of R and T is taken as R^A, we may apply Theorem 13. Hence the w_i can be taken as linearly independent and $n = (R:R^A) = (A:k) = (S:R)$. Since $(R \times_k A^*:R) = n$ and $R \times_k A^* \subset S$, we get $R \times_k A^* = S$, which means that all our homomorphisms are analytic.

Conversely, if T is simple, we can take the w's linearly independent. If we assume that all our homomorphisms are analytic, so that $S = R \times_k A^*$, we get first $n = (R \times_k A^* : R) = (A : k)$. We next get that S is simple, because it consists of all homomorphisms of the n-dimensional T-space R, and is therefore isomorphic to the set of all $n \times n$ matrices with elements in T, or $T \times_k M_n$. So $R \times_k A^*$ is simple, which implies that A is simple. Since R^A must be of degree n under R, and certainly contains T, we find that $T = R^A$, which finishes the proof.

COROLLARY. *If A is a simple subalgebra of R containing the centrum k of R, (R simple with minimal condition) and $\sigma(x)$ is a homomorphism of R satisfying the equations $\sigma(xt) = \sigma(x)t$ and $\sigma(tx) = t\sigma(x)$, for all elements t of $T = R^A$, then any expression of $\sigma(x)$ with linearly independent coefficients has the form $\Sigma a_\nu x a'_\nu$, where the a_ν as well as the a'_ν are in A.*

This is proved by the method used in the first part of the proof of the preceding theorem.

THEOREM 17. *Under the same assumptions as in the preceding corollary, every automorphism of R that leaves every element of $T = R^A$ fixed consists of a transformation by an element of A.*

Proof. Since $\sigma(xt) = \sigma(x)\sigma(t) = \sigma(x)t$ and $\sigma(tx) = t\sigma(x)$, our corollary gives $\sigma(x) = \Sigma a_\nu x a'_\nu$. We can assume that the a_ν and the a'_ν are linearly independent with respect to k. Let $l(x) = \Sigma b_\mu x b'_\mu$ be an analytic linear function of A which takes a_1 into 1 and the other a_i into 0. Compute $\sum_\mu b_\mu \sigma(b'_\mu x)$. On the one hand, it is $\{\sum_\mu b_\mu \sigma(b'_\mu)\}\sigma(x)$, and therefore is of the form $a\sigma(x)$. On the other hand, it is $\sum_{\nu,\mu} b_\mu a_\nu b'_\mu x a'_\nu = \sum_\nu l(a_\nu) x a'_\nu = x a'_1$. Here a'_1 is not 0 because it is one of a set of linearly independent elements. We therefore get $a\sigma(x) = x a'_1$. Let $\lambda(x) = \sum c_\nu x c'_\nu$ be such that $\lambda(a'_1) = 1$. Compute $\sum_\nu a\sigma(x c_\nu) c'_\nu$. On one hand, it has the form $a\sigma(x)b$; on the other hand, it is $\sum x c_\nu a'_1 c'_\nu = x$. Hence $a\sigma(x)b = x$. Putting $x = 1$, we see that $b = a^{-1}$, and hence $\sigma(x) = a^{-1}xa$. That a belongs to A follows from the corollary.

INDIANA UNIVERSITY,
INSTITUTE FOR ADVANCED STUDY.

ON WEAKLY COMPACT SUBSETS OF A BANACH SPACE.*

R. S. PHILLIPS.

In the last few years use of the Banach space weak topology has proven exceedingly fruitful in the hands of Bourbaki (4), Smulian (24), Kakutani (14), Ghantmacher and Smulian (9), Sirvint (23), Dunford and Pettis (8), and others. These investigations have been concerned either with sequentially compact subsets or with bicompact subsets. In this paper an attempt has been made to bridge the gap between these two extremes by considering $\aleph_a$-compact subsets where $\aleph_a$ is any cardinal number.

One could define a subset to be $\aleph_a$-compact if every covering of power $\leq \aleph_a$ of the set's weak closure contained a finite subcovering. Bourgin has done this, treating the problem from a topological point of view (5). We have preferred an analysis approach, which, we believe, permits a much more direct and essentially simpler development.

A subset G of a Banach space X will be said to be $\aleph_a$-compact if for every subset $[x_\pi \mid \pi \in$ directed set of power $\leq \aleph_a]$ [1] of G there exists an $x_0 \in X$ such that

$$\varlimsup_{\pi} \bar{x}(x_\pi) \leq \bar{x}(x_0) \leq \overline{\lim_{\pi}}\, \bar{x}(x_\pi)$$

for all $\bar{x}$ belonging to the adjoint space $\bar{X}$. The notion of limit is defined in the Moore-Smith sense (18). We do not insist that the limit inferior or limit superior be finite. If $\lim_{\pi} \bar{x}(x_\pi)$ exists for all $\bar{x} \in \bar{X}$, it is clear that $\bar{x}(x_0) = \lim_{\pi} \bar{x}(x_\pi)$. A subset G will be called $\aleph_a$-*complete* if for every subset $[x_\pi \mid \pi \in$ directed set of power $\leq \aleph_a]$ of G for which $\lim_{\pi} \bar{x}(x_\pi)$ exists for all $\bar{x} \in \bar{X}$, there exists an x_0 such that $\bar{x}(x_0) = \lim_{\pi} \bar{x}(x_\pi)$ for all $\bar{x} \in \bar{X}$. Such an x_0 is necessarily unique. If, then, G is $\aleph_a$-compact, it is likewise $\aleph_a$-complete. The converse is not true. An $\aleph$-*compact* subset G will be defined as a set which is $\aleph_a$-compact for all $\aleph_a$.

Many of our results are simply generalizations of known theorems for sequentially weak compact sets. For purposes of completeness we have given new proofs for some of these theorems. We have tried to free ourselves from the

* Received August 29, 1941.

[1] The class of elements s satisfying the property P will be designated by $[s \mid P]$.

analogues of the usual separability condition which has often accompanied the discussion of sequentially weak compact sets. This we have been able to do in the statement of our results, but unfortunately not always in our method of proof.

In Section 1 we have compared $\aleph_a$-compactness with related compactness notions. We find, for instance, that for weakly closed sets coverings of power $\leq \aleph_a$ have finite subcoverings if and only if the sets are $\aleph_a$-compact, while for convex sets $\aleph_a$-compactness is equivalent to transfinite closedness of order $\aleph_a$. If the unit sphere of X is $\aleph_a$-compact, then the unit sphere of $\bar{X}$ is also $\aleph_a$-compact. If in addition there exists a set of power $\leq \aleph_a$ total on X, then X is reflexive.

Section 2 deals with the obvious generalization of the weakly completely continuous transformations, namely, $\aleph_a$-compact transformations. If U is $\aleph_a$-compact, then the adjoint transformation $\bar{U}$ is likewise $\aleph_a$-compact. If a sequence of $\aleph_a$-compact transformations converges in norm to a transformation U, then U is found to be $\aleph_a$-compact.

By the principal theorem of Section 3, the closure of the set $[\Sigma_G a_x x \mid \Sigma_G \mid a_x \mid \leq 1]$ is $\aleph_a$-compact if and only if G is $\aleph_a$-compact. Hence the closed convex extension of an $\aleph_a$-compact set is $\aleph_a$-compact. An application of this theorem gives us a criterion for $\aleph_a$-compact sets.

Section 4 is concerned with determining conditions on the geometry of X under which $\aleph_a$-compactness implies $\aleph$-compactness. For example, all $\aleph_0$-compact subsets of a strictly convex space are $\aleph$-compact, and all $\aleph_a$-compact subsets of a space for which there exists a total set of power $\leq \aleph_a$ are $\aleph$-compact. We have as yet not been able to find an example of an $\aleph_0$-compact set which is not $\aleph$-compact.

An application of our results is given in the final section of the paper. Here we prove that if $x(\tau)$ is a completely additive, absolutely continuous set function of bounded variation such that $[(x(\tau)/|\tau|) \mid (||x(\tau)||/|\tau|) \leq n]$ is $\aleph_0$-compact for all n, then there exists a measurable function $x(t)$ such that $x(\tau) = \int_\tau x(t)\,d\alpha$. This leads to a general form of the weakly completely continuous transformation on $\mathfrak{L}$ to an arbitrary Banach space and to the conclusion that this transformation is necessarily separable.

It will be convenient to make the following notational conventions: The weak topology $w(X)$ of X as elements has the generic neighborhood $U(x_0) \equiv [x \mid\mid \bar{x}_i(x - x_0)| < \epsilon \quad i = 1, 2, \cdots, n]$ where the $\bar{x}_i$ are arbitrary elements of $\bar{X}$. The weak topology $\bar{w}(\bar{X})$ of $\bar{X}$ as functionals has the generic neighhorhood $U(\bar{x}_0) \equiv [\bar{x} \mid\mid (\bar{x} - \bar{x}_0)x_i \mid < \epsilon \quad i = 1, 2, \cdots, n]$ where the x_i are

arbitrary elements of X. X is a regular Hausdorf space in the $w(X)$ topology; while $\bar{X}$ is a bicompact Hausdorf space in the $w(\bar{X})$ topology. If G is a subset of X, then G_L will designate the smallest linear extension of G; G_C will be the norm closure of G; and $cv(G)$ will be the convex extension of G. M_T will be the Banach space of bounded functions $a(t)$ on an abstract class $T \equiv [t]$ to real numbers having the norm $\| a \| = \text{l. u. b. } [| a(t) | \mid t \epsilon T]$. l_T will be the Banach space of functions $a(t)$ on an abstract class $T \equiv [t]$ to real numbers for which $\| a \| = \sum_T | a(t) |$ exists. If $f(t)$ is a function, then $f(\cdot)$ will designate the function considered as an element of some class of functions. Finally, $X_C \equiv [x \mid x \epsilon X, \| x \| \leq c]$ so that X_1 is the unit sphere.

1. Related compactness definitions. Before attempting a systematic study of the properties of $\aleph_a$-compact sets, we shall illustrate the usefulness of this notion by giving a criterion for reflexive Banach spaces. A Banach space is called *reflexive* if for every $\bar{\bar{x}} \epsilon \bar{\bar{X}}$, there exists an $x \epsilon X$ such that $\bar{\bar{x}}(\bar{x}) = \bar{x}(x)$ for all $\bar{x} \epsilon \bar{X}$.

1. 1. THEOREM. *A necessary and sufficient condition that X be reflexive is that X_1 be $\aleph_a$-compact and that $\bar{X}$ contain a subset of power $\leq \aleph_a$ total on X.*

Given $\bar{\bar{x}} \epsilon \bar{\bar{X}}$ and Γ, an arbitrary subset of $\bar{X}$ of power $\leq \aleph_a$, we first show that there exists an $x_\Gamma \epsilon X$ such that $\bar{\bar{x}}(\bar{x}) = \bar{x}(x_\Gamma)$ for all $\bar{x} \epsilon \Gamma$. The set of all finite subsets π of Γ is a directed set ($\leq$ means set inclusion) of power $\leq \aleph_a$. By Helly's theorem (13), there exists an $x_\pi \epsilon X$ such that $\bar{\bar{x}}(\bar{x}) = \bar{x}(x_\pi)$ for each $\bar{x} \epsilon \pi$ and $\| x_\pi \| \leq \| \bar{\bar{x}} \| + 1$. As X_1 (and hence X_n) is $\aleph_a$-compact, there exists an $x_\Gamma \epsilon X$ satisfying the condition

$$\underline{\lim}_\pi \bar{x}(x_\pi) \leq x(x_\Gamma) \leq \overline{\lim}_\pi \bar{x}(x_\pi)$$

for all $\bar{x} \epsilon \bar{X}$. Now $\lim_\pi \bar{x}(x_\pi) = \bar{\bar{x}}(\bar{x})$ for all $\bar{x} \epsilon \Gamma$. Therefore $\bar{\bar{x}}(\bar{x}) = \bar{x}(x_\Gamma)$ for all $\bar{x} \epsilon \Gamma$.

Let us now apply this first to the case where Γ is a set of power $\leq \aleph_a$ total on X, and then to $\Gamma' = \Gamma \cup \bar{x}_0$ where $\bar{x}_0$ is an arbitrary element of $\bar{X}$. Then $\bar{x}(x_\Gamma) = \bar{\bar{x}}(\bar{x}) = \bar{x}(x_{\Gamma'})$ for all $\bar{x} \epsilon \Gamma$. As Γ is total, $x_\Gamma = x_{\Gamma'}$. Thus $\bar{\bar{x}}(\bar{x}) = \bar{x}(x_\Gamma)$ for all $\bar{x}$ belonging to Γ' and hence, since $\bar{x}_0$ is arbitrary, for all $\bar{x} \epsilon \bar{X}$. In other words, X is reflexive.

The necessity is quite evident. For, given any set $[x_\pi] \subset X_1$ where $[\pi]$ is a directed set, an argument similar to that used by Banach (1, p. 118) demonstrates the existence of an $\bar{\bar{x}} \epsilon \bar{\bar{X}}$ such that

$$\underline{\lim}_\pi \bar{x}(x_\pi) \leq \bar{\bar{x}}(\bar{x}) \leq \overline{\lim}_\pi \bar{x}(x_\pi)$$

for all $\bar{x} \in \bar{X}$. As X is reflexive we can replace $\bar{x}$ by an element of X. As this holds for all sets $[\pi]$, it holds in particular for all sets of power less than or equal to the power of any subset of $\bar{X}$ total on X.

1.2 *Definition.* A set $G \subset X$ is said to be $\aleph_a$-*compact** if for any family of subsets $[A_t \mid t \in T]$ of G such that 1). A_t is closed in the $w(X)$ topology; 2). the power of $T \leq \aleph_a$; 3). any finite subset has a non-vacuous intersection; then $\bigcap_T A_t$ is non-vacuous.

D. G. Bourgin (5) has shown that if X_1 is $\aleph_a$-compact* and if $\bar{X}$ contains a subset of power $\leq \aleph_a$ which is $\bar{w}(\bar{X})$ dense, then X is reflexive. In Theorem 1.9 we show that X_1 is $\aleph_a$-compact if and only if it is $\aleph_a$-compact*. Further if $\Gamma \subset \bar{X}$ is total on X then the set of all finite linear combinations with rational coefficients, which is of the same power, is $\bar{w}(\bar{X})$ dense in $\bar{X}$ (see Lemma 2.4). Hence the two theorems are equivalent.

In the statement of Theorem 1.1, the hypothesis that X be $\aleph_a$-compact can not be replaced by $\aleph_a$-completeness in the $w(X)$ topology. This breaks down for the space l which is $\aleph_0$-complete in the $w(l)$ topology and for which any dense subset of the separable space c_0 is total.

For purposes of orientation, we shall next compare $\aleph_a$-compactness with related compactness notions. In this it will be convenient to have the following lemmas.

1.3 LEMMA. *If G is $\aleph_a$-compact, then given $[x_\pi \mid \pi \in$ directed set of power $\leq \aleph_a] \subset G$, there exists an x_0 belonging to the closed convex extension of G such that*

$$\underline{\lim_\pi}\, \bar{x}(x_\pi) \leq \bar{x}(x_0) \leq \overline{\lim_\pi}\, \bar{x}(x_\pi)$$

for every $\bar{x} \in \bar{X}$.

If x_0 did not belong to the closed convex extension of G, there would exist an $\bar{x}_0$ and a constant c such that $\bar{x}_0(x_0) > c$ while $\bar{x}_0(x) \leq c$ for all $x \in G$ (8, p. 331). This clearly contradicts the inequalities satisfied by x_0.

1.4 LEMMA. *If G is $\aleph_a$-compact, $[x_\pi \mid \pi \in$ directed set of power $\leq \aleph_a] \subset G$, and if $\lim_\pi \bar{x}(x_\pi)$ exists for a set $\Gamma \subset \bar{X}$ total on G_{LC}, then x_π converges weakly to a unique element of X.*

As G is $\aleph_a$-compact, there exists by Lemma 1.3 an $x_0 \in G_{LC}$ such that

$$\underline{\lim_\pi}\, \bar{x}(x_\pi) \leq \bar{x}(x_0) \leq \overline{\lim_\pi}\, \bar{x}(x_\pi)$$

for all $\bar{x} \in \bar{X}$. Suppose that for a given $\bar{x}_0 \in \bar{X}$, $\lim_\pi \bar{x}_0(x_\pi) \neq \bar{x}_0(x_0)$. Then

there exists an $\epsilon > 0$ and a subset $[\pi']$ of $[\pi]$ cofinal with $[\pi]$ such that $\underline{\lim}_{\pi'} \bar{x}_0(x_{\pi'}) \geq \bar{x}_0(x_0) + \epsilon$ (or $\overline{\lim}_{\pi'} \bar{x}_0(x_{\pi'}) \leq \bar{x}(x_0) - \epsilon$). As before there exists an $x'_0 \epsilon G_{LC}$ such that

$$\underline{\lim}_{\pi'} \bar{x}(x_{\pi'}) \leq \bar{x}(x'_0) \leq \overline{\lim}_{\pi'} \bar{x}(x_{\pi'})$$

for all $\bar{x} \epsilon \bar{X}$. Now $\bar{x}(x_0) = \bar{x}(x'_0)$ for all $\bar{x} \epsilon \Gamma$. Therefore $x_0 = x'_0$, contrary to $\bar{x}_0(x'_0) \geq \bar{x}_0(x\ ) + \epsilon$. Hence $\lim_{\pi} \bar{x}(x_\pi) = \bar{x}(x_0)$ for all $\bar{x} \epsilon \bar{X}$. x_0 is clearly unique.

1.5 *Definition.* G is said to be $\aleph_a$-*closed* if, given any subset $[x_\mu]$ where $[\mu]$ is a *well-ordered* set of power $\leq \aleph_a$, there is an $x_0 \epsilon X$ such that

$$\underline{\lim}_{\mu} \bar{x}(x_\mu) \leq \bar{x}(x_0) \leq \overline{\lim}_{\mu} \bar{x}(x_\mu)$$

for all $\bar{x} \epsilon \bar{X}$.

1.6 THEOREM. *The following are equivalent:*

1). *G is $\aleph_0$-compact,*

2). *G is sequentially compact in the $w(X)$ topology,*

3). *G is $\aleph_0$-closed.*

Let $\{x_n\}$ be a subset of G. Then $X_0 = \{x_n\}_{LC}$ is separable. We can therefore choose a denumerable set of functionals on X_0, dense in the $\bar{w}(\bar{X}_0)$ topology (1, pp. 185-186). These functionals can be extended to be functionals on X (1, p. 55). The set of extended functionals, Γ, will clearly be total on X_0. Using the diagonal process, we can obtain a subsequence $\{x_{n'}\}$ for which $\lim \bar{x}(x_{n'})$ exists for all $\bar{x} \epsilon \Gamma$. As $G \cap X_0$ is again $\aleph_0$-compact (Lemma 1.3) we can apply Lemma 1.4. Thus $\{x_{n'}\}$ converges weakly to an element of X. Hence 1) implies 2). 2) clearly implies 3). To prove that 3) implies 1), suppose we are given a denumerable subset $\{x_\pi\}$ of G. Let us order the π's $(\pi_1, \pi_2, \cdots)$ and choose $\pi^n \geq (\pi_1, \pi_2, \cdots, \pi_n)$. The set (π^n) is clearly cofinal with (π). By hypothesis there exists an $x_0 \epsilon X$ such that $\underline{\lim}_{n} \bar{x}(x_{\pi^n}) \leq \bar{x}(x_0) \leq \overline{\lim}_{n} \bar{x}(x_{\pi^n})$ for all $\bar{x} \epsilon \bar{X}$. It follows that

$$\underline{\lim}_{\pi} \bar{x}(x_\pi) \leq \bar{x}(x_0) \leq \overline{\lim}_{\pi} \bar{x}(x_\pi)$$

for all $\bar{x} \epsilon \bar{X}$.

1.7 COROLLARY. *If G is $\aleph_0$-compact, then G is bounded.*

1.8 LEMMA. *Given a directed set $\Pi = [\pi]$, let S_α, for each $\alpha \epsilon A$, be a*

bicompact metric space containing the set $[a_{\alpha\pi} \mid \pi \in \Pi]$. *Then there exists a directed set* $\Sigma = [\sigma]$ *and a function* π_σ *on* Σ *to* Π *for which*

1. $\lim_\sigma a_{\alpha\pi_\sigma}$ *exists in* S_α.

2. *Given* π_0, *there exists a* σ_0 *such that* $\pi_\sigma \geq \pi_0$ *for all* $\sigma \geq \sigma_0$.

3. *The power of* Σ *is less than or equal to the power of* $A + \Pi$.

Let P_A be the product space of the spaces S_α where the generic point is $p = [p(\alpha) \in S_\alpha]$ and the generic neighborhood is $U(p_0) \equiv [p \mid \rho[p_0(\alpha_i), p(\alpha_i)] < \epsilon \quad i = 1, 2, \cdots, k]$. Čech (6, p. 830) has shown that this is a bicompact Hausdorff space. Define $p_\pi = [a_{\alpha\pi} \mid \alpha \in A] \in P_A$. Let $F_{\pi_0} \equiv [p_\pi \mid \pi \geq \pi_0]$. Since Π is a directed set, all finite sets of F_π's have a non-vacuous intersection. Therefore there exists a $p \in \bigcap \overline{F_\pi}$. We define Σ to consist of all combinations $[\pi, \alpha_1, \alpha_2, \cdots, \alpha_k]$ where $\pi \in \Pi$ and $\alpha \in A$. $\sigma_1 \geq \sigma_2$ will mean that the α's in σ_2 are a subset of those in σ_1 and $\pi_1 \geq \pi_2$. We define the function π_σ as follows: If $\sigma = [\pi_0, \alpha_1, \cdots, \alpha_k]$, then as $p \in \overline{F_{\pi_0}}$ there will be a $p_{\pi_1} \in F_{\pi_0}$ such that $\rho[p_{\pi_1}(\alpha_i), p(\alpha_i)] < 1/k$, $i = 1, 2, \cdots, k$, $\pi_1 \geq \pi_0$. Define $\pi_\sigma = \pi_1$. Clearly $\lim_\sigma p_{\pi_\sigma} = p$. Conditions 2) and 3) are likewise satisfied.

1.9 THEOREM. *A set* G, *closed in the* $w(X)$ *topology is* $\aleph_a$*-compact if and only if it is* $\aleph_a$*-compact*.*

Suppose first that G is $\aleph_a$-compact and that $[A_t \mid t \in T]$ is a family of $w(X)$ closed subsets of G such that any finite set has a non-vacuous intersection and T is of power $\leq \aleph_a$. Let $[\pi]$ be a directed family of finite subsets of T. The power of $[\pi]$ is then less than or equal to $\aleph_a$. Choose $x_\pi \in \bigcap_{t \in \pi} A_t$. Let H be the set of all finite linear combinations of the x_π with rational coefficients and let $X_0 = H_G = [x_\pi]_{LC}$. Clearly H is of power $\leq \aleph_a$. For every $x \in H$, choose an $\bar{x} \in \bar{X}$ such that $\bar{x}(x) \geq ||x||/2$ and $||\bar{x}|| = 1$. Then the set Γ of $\bar{x}$'s so chosen will be total on X_0 and of power $\leq \aleph_a$. Applying Lemma 1.8 to the set of numbers $[\bar{x}(x_\pi) \mid \bar{x} \in \Gamma, \pi \in [\pi]]$ we obtain a directed set $[\sigma]$ of power $\leq \aleph_a$ for which $\lim_\sigma \bar{x}(x_{\pi_\sigma})$ exists for all $\bar{x} \in \Gamma$. By Lemma 1.4, there exists an $x_0 \in X$ such that $\lim_\sigma \bar{x}(x_{\pi_\sigma}) = \bar{x}(x_0)$ for all $\bar{x} \in \bar{X}$. Now by Lemma 1.8, given $\pi_0 \equiv t_0$, there exists a σ_0 such that $\pi_\sigma \geq \pi_0$ for all $\sigma \geq \sigma_0$. In other words, $x_{\pi_\sigma} \in A_{t_0}$ for all $\sigma \geq \sigma_0$. As A_t is $w(X)$ closed, x_0 belongs to A_t, i.e., $x_0 \in \bigcap_{t \in T} A_t$.

To prove the converse, suppose G to be $\aleph_a$-compact*. Let $[x_\pi] \subset G$ where $[\pi]$ is a directed family of power $\leq \aleph_a$. Define $F_{\pi_0} \equiv [x_\pi \mid \pi \geq \pi_0]$ and $A_\pi \equiv w(X)$ closure of F_π. As $[\pi]$ is a directed set, any finite set of F_π's (and

hence A_π's) have a non-vacuous intersection. There will therefore exist an $x_0 \in \bigcap A_\pi$. Finally given π_0, $\bar{x}_0$, and $\epsilon > 0$ there exists an $x_\pi \in F_{\pi_0}$ for which $|\bar{x}_0(x_\pi - x_0)| < \epsilon$. As $x_\pi \in F_{\pi_0}$ implies $\pi \geq \pi_0$, we have

$$\underline{\lim}_\pi \bar{x}(x_\pi) \leq \bar{x}(x_0) \leq \overline{\lim}_\pi \bar{x}(x_\pi)$$

for all $\bar{x} \in \bar{X}$.

1.10 THEOREM. *A convex set G is $\aleph_a$-compact if and only if it is $\aleph_a$-closed.*

As any well-ordered set is a directed set, it is clear that $\aleph_a$-compact implies $\aleph_a$-closed. Suppose that G is convex and $\aleph_a$-closed. By the remark following Theorem 3.4, the norm closure of G is likewise $\aleph_a$-closed. For convex sets the norm closure and the $w(X)$ closure are identical (15). We may therefore suppose that G is a convex $w(X)$ closed set. Let $[x_\pi \mid \pi \in \text{directed set of power } \leq \aleph_a] \subset G$ and define F_{π_0} to be the convex extension of $[x_\pi \mid \pi \geq \pi_0]$. Let A_π be the $w(X)$ closure of F_π. It is clear that all finite sets of A_π's have a non-vacuous intersection. Assume that this is true for all cardinals less than $\aleph_b \leq \aleph_a$. Then given any set of A_π's of power $\aleph_b$, we well-order this set with ordinals μ less that the first ordinal O_b having $\aleph_b$ predecessors. For each $\lambda < O_b$, choose an $x_\lambda \in \bigcap_{\mu<\lambda} A_\mu$. As G is $\aleph_a$-closed there exists an x_0 for which

$$\underline{\lim}_{\lambda \to O_b} \bar{x}(x_\lambda) \leq \bar{x}(x_0) \leq \overline{\lim}_{\lambda \to O_b} \bar{x}(x_\lambda)$$

for all $\bar{x} \in \bar{X}$. If x_0 were not an element of A_{π_λ}, then we could find an $\bar{x} \in \bar{X}$ and a constant c such that $\bar{x}(A_{\pi_\lambda}) \leq c$ and $\bar{x}(x_0) > c$ (8, p. 331). Thus $\bar{x}(x_\mu) \leq c$ for all $\mu \geq \lambda$, contrary to the above inequalities. This establishes the induction. Hence let x belong to $\bigcap A_\pi$. Then for a given ϵ, $\bar{x}$, π_0 there exist $\pi_1, \pi_2, \cdots, \pi_n \geq \pi_0$ such that $|\bar{x}(\sum_{i=1}^n c_i x_{\pi_i} - x)| < \epsilon$ where $c_i \geq 0$ and $\sum_{i=1}^n c_i = 1$. For one of these π_i's, $\bar{x}(x_\pi) \leq \bar{x}(x) + \epsilon$ while for another $\bar{x}(x_\pi) \geq \bar{x}(x) - \epsilon$. Hence

$$\underline{\lim}_\pi \bar{x}(x_\pi) \leq \bar{x}(x) \leq \overline{\lim}_\pi \bar{x}(x_\pi).$$

1.11 *Definition.* A space X is said to satisfy the *Helly property of order $\aleph_a$*, if whenever

$$|\sum_{i=1}^n a_i c_{a_i}| \leq M \,||\sum_{i=1}^n a_i \bar{x}_{a_i}||,$$

where the a_i are arbitrary real numbers, then the system $\bar{x}_a(x) = c_a$ consisting of at most $\aleph_a$ equations has a solution $x \in X_M$.

1.12 LEMMA. *If X_1 is $\aleph_a$-compact, then X satisfies the Helly property of order $\aleph_a$.*

Let us suppose the conditions stated in Definition 1.11 are satisfied by the functionals $[\bar{x}_\alpha]$ and the numbers $[c_\alpha]$. Let π be a finite subset $(\alpha_1, \alpha_2, \cdots, \alpha_k)$ of $[\alpha]$. Then $[\pi]$ is a directed set ($\leq$ means set inclusion) of power $\leq \aleph_a$. By Helly's theorem (14), we can find for each π an x_π such that $\bar{x}_\alpha(x_\pi) = c_\alpha$ for every $\alpha \in \pi$ and $\| x_\pi \| \leq M + 1/k$. As X_{M+1} is $\aleph_a$-closed there exists an x_0 such that

$$\underline{\lim_\pi}\, \bar{x}(x_\pi) \leq \bar{x}(x_0) \leq \overline{\lim_\pi}\, \bar{x}(x_\pi)$$

for each $\bar{x} \in \bar{X}$. Since $\lim_\pi \bar{x}_\alpha(x_\pi)$ exists, we have $\bar{x}_\alpha(x_0) = c_\alpha$. Finally $|\bar{x}(x_0)| \leq \overline{\lim_\pi} |\bar{x}(x_\pi)| \leq M \|\bar{x}\|$ implies $\| x_0 \| \leq M$.

1.13 LEMMA. *If X satisfies the Helly property of order $\aleph_a$, then $\bar{X}_1$ is $\aleph_a$-compact.*

Let $[\bar{x}_\pi \mid \pi \in$ directed set of power $\leq \aleph_a]$ be bounded in the norm by one. By an argument similar to that used by Banach (1, p. 118) we can obtain an $\bar{x}_0 \in \bar{X}$ for which

$$\underline{\lim_\pi}\, \bar{x}_\pi(x) \leq \bar{x}_0(x) \leq \overline{\lim_\pi}\, \bar{x}_\pi(x)$$

for all $x \in X$. Now if $\bar{\bar{x}} \in \bar{\bar{X}}$, then

$$\left| \sum_{i=1}^{n} a_i \bar{\bar{x}}(\bar{x}_{\alpha_i}) \right| \leq \|\bar{\bar{x}}\| \cdot \left\| \sum_{i=1}^{n} a_i \bar{x}_{\alpha_i} \right\|$$

where $\bar{x}_\alpha \in [\bar{x}_\pi]$ or $\bar{x}_\alpha = \bar{x}_0$. Hence the Helly property of order $\aleph_a$ implies the existence of an x_0 for which $\bar{x}_\alpha(x_0) = \bar{\bar{x}}(\bar{x}_\alpha)$ for all π and $\bar{x}_0(x_0) = \bar{\bar{x}}(\bar{x}_0)$. Consequently

$$\underline{\lim_\pi}\, \bar{\bar{x}}(\bar{x}_\pi) = \underline{\lim_\pi}\, \bar{x}_\pi(x_0) \leq \bar{x}_0(x_0) = \bar{\bar{x}}(\bar{x}_0) \leq \overline{\lim_\pi}\, \bar{x}_\pi(x_0) = \underline{\lim_\pi}\, \bar{\bar{x}}(\bar{x}_\pi),$$

so that $\bar{X}_1$ is $\aleph_a$-compact.

1.14 THEOREM. *The following statements are equivalent:*

1). *X_1 is $\aleph_a$-compact.*

2). *X satisfies the Helly property of order $\aleph_a$.*

3). *$\bar{X}_1$ is $\aleph_a$-compact.*

Lemmas 1.12 and 1.13 show that 1) implies 2) implies 3). Reiterating this reasoning we find that 3) implies that $\bar{X}_1$ is $\aleph_{1a}$-compact. As X_1 is a closed convex subset of $\bar{X}_1$, it follows that $\bar{X}_1$ is $\aleph_a$-compact (see Lemma 1.3).

A property will be said to be *hereditary* if whenever it holds for a Banach space, it also holds for every closed linear subspace. Theorem 1.14 and Lemma 1.3 give

1.15 COROLLARY. *The Helly property of order $\aleph_a$ is hereditary.*

2. On weakly completely continuous transformations. We shall consider in this section a generalization of the weakly completely continuous transformations of a Banach space X to another Banach space Y.

Definition. A linear transformation U on X to Y will be called *$\aleph_a$-compact* if $U(X_1)$ is $\aleph_a$-compact.

This definition is an extension of the usual notion of a completely continuous transformation. In view of the following lemma, it would clearly not do to require only that the transformation take $\aleph_a$-compact sets into $\aleph_a$-compact sets.

2.1 LEMMA. *If G is an $\aleph_a$-compact subset of X and U is a linear limited transformation on X to Y, then $U(G)$ is $\aleph_a$-compact.*

Suppose $[y_\pi \mid \pi \in \text{directed set of power} \leq \aleph_a] \subset U(G)$. Then y_π is the image of some $x_\pi \in G$. G being $\aleph_a$-compact, there exists an $x_0 \in X$ such that

$$\varliminf_\pi \bar{x}(x_\pi) \leq \bar{x}(x_0) \leq \varlimsup_\pi \bar{x}(x_\pi)$$

for all $\bar{x} \in \bar{X}$. Consequently

$$\varliminf_\pi \bar{y}(y_\pi) = \varliminf_\pi \bar{U}\bar{y}(x_\pi) \leq \bar{U}\bar{y}(x_0) = \bar{y}(Ux_0) \leq \varlimsup_\pi \bar{U}\bar{y}(x_\pi) = \varlimsup_\pi \bar{y}(y_\pi)$$

for all $\bar{y} \in \bar{Y}$.

2.2 THEOREM. *If U is $\aleph_a$-compact, then the adjoint transformation $\bar{U}$ is $\aleph_a$-compact.*

Let $[\bar{y}_\pi \mid \pi \in \text{directed set of power} \leq \aleph_a]$ be bounded by one. Then there exists a $\bar{y}_0 \in \bar{Y}$ such that

$$\varliminf_\pi \bar{y}_\pi(y) \leq \bar{y}_0(y) \leq \varlimsup_\pi \bar{y}_\pi(y)$$

for all $y \in Y$. We shall show that for any $\bar{\bar{x}} \in \bar{\bar{X}}$

$$\varliminf_\pi \bar{\bar{x}}\bar{U}(\bar{y}_\pi) \leq \bar{\bar{x}}\bar{U}(\bar{y}_0) \leq \varlimsup_\pi \bar{\bar{x}}\bar{U}(\bar{y}_\pi).$$

Define σ to be a finite subset of $[\bar{y}_\pi] \cup \bar{y}_0$. Then $[\sigma]$ is a directed set ($\leq$ means set inclusion) of power $\leq \aleph_a$. By Helly's theorem (13), for each σ there exists an $x_\sigma \in X$ such that both $\bar{U}(\bar{y})x_\sigma = \bar{\bar{x}}\bar{U}(\bar{y})$ for $\bar{y} \in \sigma$ and $\| x_\sigma \| \leq \| \bar{\bar{x}} \| + 1$. Hence $\lim_\sigma \bar{y}U(x_\sigma) = \bar{\bar{x}}\bar{U}(\bar{y})$ exists for any $\bar{y}$ of the

set $[\bar{y}_\pi] \cup \bar{y}_0$. As $U(X_1)$ is $\aleph_a$-compact, there exists a $y_0 \epsilon Y$ for which

$$\underline{\lim_\sigma}\, \bar{y}U(x_\sigma) \leq \bar{y}(y_0) \leq \overline{\lim_\sigma}\, \bar{y}U(x_\sigma)$$

for all $\bar{y} \epsilon \bar{Y}$. Clearly $\bar{\bar{x}}\bar{U}(\bar{y}) = \lim_\sigma \bar{y}U(x_\sigma) = \bar{y}(y_0)$ for all $\bar{y}$ of the set $[\bar{y}_\pi] \cup \bar{y}_0$. Consequently

$$\underline{\lim_\pi}\, \bar{\bar{x}}\bar{U}(\bar{y}_\pi) = \underline{\lim_\pi}\, \bar{y}_\pi(y_0) \leq \bar{y}_0(y_0) = \bar{\bar{x}}\bar{U}(\bar{y}_0) \leq \overline{\lim_\pi}\, \bar{y}_\pi(y_0) = \overline{\lim_\pi}\, \bar{\bar{x}}\bar{U}(\bar{y}_\pi).$$

2.3 COROLLARY. *If $\bar{U}$ is $\aleph_a$-compact, then U is likewise $\aleph_a$-compact.*

By Theorem 2.2, if $\bar{U}$ is $\aleph$-compact, then $\bar{\bar{U}}$ is $\aleph_a$-compact. That is, $\bar{\bar{U}}(\bar{\bar{X}}_1)$ is $\aleph_a$-compact. But X_1 is equivalent to a subset G of $\bar{\bar{X}}_1$. Therefore $U(X_1) \equiv \bar{\bar{U}}(G)$ is $\aleph_a$-compact.

Theorem 2.2 is a generalization of a theorem due to V. Gantmacher (10).

The following lemmas will be found useful in our further work on $\aleph_a$-compact transformations. Lemma 2.4, which is essentially a simplification of a lemma due to Banach (1, pp. 119-121), was first stated by Kakutani (14).

2.4 LEMMA. *If Γ is a $\bar{w}(\bar{X})$ closed linear subset of $\bar{X}$ not containing $\bar{x}_0$, then there exists an $x_0 \epsilon X$ such that $\bar{x}_0(x_0) \neq 0$ and $\bar{x}(x_0) = 0$ for all $\bar{x} \epsilon \Gamma$.*

As $\bar{X}$ is a regular Hausdorf space in the $\bar{w}(\bar{X})$ topology, there exists a neighborhood $U_{x_0} \equiv [\bar{x} \,|\, |(\bar{x} - \bar{x}_0)x_i| < \epsilon,\ i = 1, 2, \cdots, n]$ containing no elements of Γ. Consider the linear transformation $T(\bar{x}) = \{\bar{x}(x_1), \bar{x}(x_2), \cdots, \bar{x}(x_n)\}$ which maps $\bar{X}$ into the n-dimensional euclidean space E. $T(\Gamma)$ is a linear subspace of E not containing $T(\bar{x}_0)$. Hence there exists an $(n-1)$-dimensional hyperplane $\{c_i\}$ containing $T(\Gamma)$ but not containing $T(\bar{x}_0)$. That is

$$\bar{x}\Big(\sum_{i=1}^{n} c_i x_i\Big) = \sum_{i=1}^{n} c_i \bar{x}(x_i) \begin{cases} = 0 & \text{for all } \bar{x} \epsilon \Gamma \\ \neq 0 & \text{for } \bar{x} = \bar{x}_0 \end{cases}$$

$x_0 = \sum_{i=1}^{n} c_i x_i$ satisfies the conditions of the lemma.

2.5 LEMMA. *If G is $\aleph_a$-compact and if there exists a set $\Gamma \subset \bar{X}$ of power $\leq \aleph_a$ total on G_{LC}, then G is $\aleph$-compact.*

Suppose $[x_\pi \mid \pi \epsilon \text{ directed set}] \subset G$. Applying Lemma 1.8 to the set $[\bar{x}(x_\pi) \mid \bar{x} \epsilon \bar{X}, \pi \epsilon [\pi]]$, we obtain a directed set $[\sigma]$ such that $\lim_\sigma \bar{x}(x_{\pi_\sigma})$ exists for all $\bar{x} \epsilon \bar{X}$. The set of all finite subsets ν of Γ is a directed set ($\leq$ means set inclusion) of power $\leq \aleph_a$. For $\nu = (\bar{x}_1, \bar{x}_2, \cdots, \bar{x}_k)$ there exists an x_{π_σ} such that $|\bar{x}(x_{\pi_\sigma}) - \lim_\sigma \bar{x}(x_{\pi_\sigma})| < 1/k$ for all $\bar{x} \epsilon \nu$. Define this x_{π_σ} to be x_ν. Then $\lim_\nu \bar{x}(x_\nu) = \lim_\sigma \bar{x}(x_{\pi_\sigma})$ for all $\bar{x} \epsilon \Gamma$. As G is $\aleph_a$-compact there exists

an x_0 such that $\lim_\nu \bar{x}(x_\nu) = \bar{x}(x_0)$ for all $\bar{x} \in \Gamma$. If we repeat this process for $\Gamma' = \Gamma \cup \bar{x}_0$, where $\bar{x}_0$ is an arbitrary element of $\bar{X}$, we obtain an x'_0 such that $\bar{x}(x'_0) = \lim_\sigma \bar{x}(x_{\pi_\sigma})$ for all $\bar{x} \in \Gamma'$ and $\bar{x}(x_0) = \bar{x}(x'_0)$ for all $\bar{x}$ belonging to the total set Γ. Hence $x_0 = x'_0$ and as $\bar{x}_0$ was arbitrary, $\bar{x}(x_0) = \lim_\sigma \bar{x}(x_{\pi_\sigma})$ for all $\bar{x} \in \bar{X}$. Now, given π_0, there exists a σ such that $\pi_\sigma \geq \pi_0$ for all $\sigma \geq \sigma_0$. It follows that

$$\underline{\lim_\pi}\, \bar{x}(x_\pi) \leq \underline{\lim_\sigma}\, \bar{x}(x_{\pi_\sigma}) = \bar{x}(x_0) = \overline{\lim_\sigma}\, \bar{x}(x_{\pi_\sigma}) \leq \overline{\lim_\pi}\, \bar{x}(x_\pi).$$

Lemma 2.5 is a rather strong result. For instance, if X_1 is $\aleph_a$-compact and there exists a subset of $\bar{X}$ of power $\leq \aleph_a$ total on X, it follows from this lemma, that X is $\aleph_a$-complete for all $\aleph$. This, by a result due to Goldstine (12), implies that X is reflexive (Theorem 1.1).

2.6 THEOREM. *If U is $\aleph_a$-compact and if there exists a set of power $\leq \aleph_a$ total on $U(X)_C$ then* (1) *U is $\aleph$-compact fo rall $\aleph$, and* (2) *$\bar{U}(\bar{Y})$ is $\aleph_a$-separable.*[2]

By Lemma 2.5 U is $\aleph$-compact for all $\aleph$. Let $Y_0 = U(X)_C$. U also defines an adjoint transformation $\bar{U}_1$ on $\bar{Y}_0$ to $\bar{X}$. Since 1) any functional on Y_0 can be extended to a functional on Y and conversely any functional on Y defines a functional on Y_0, and 2) $\bar{U}(\bar{y})$ depends only on the behavior of $\bar{y}$ on Y_0, it follows that $\bar{U}(\bar{Y}) \equiv \bar{U}_1(\bar{Y}_0)$. By hypothesis, there exists a subset of $\bar{Y}_0$ of power $\leq \aleph_a$ total on Y_0. Let Γ be the set of all finite linear combinations with rational coefficients of this set. Γ is of power $\leq \aleph_a$ and $\Gamma_C = (\Gamma_L)_C$. $\bar{Y}_0$ is the $\bar{w}(\bar{Y}_0)$ closure of Γ_L, by Lemma 2.4, and hence of Γ. Therefore given $\bar{y} \in \bar{Y}_0$, there exists a set $[\bar{y}_\pi \mid \pi \in \text{directed set}] \subset \Gamma$ such that $\lim_\pi \bar{y}_\pi(y) = \bar{y}(y)$ for all $y \in Y_0$. If we now employ the reasoning used in Theorem 2.2, we obtain $\lim_\pi \bar{\bar{x}}\bar{U}_1(\bar{y}_\pi) = \bar{\bar{x}}\bar{U}_1(\bar{y})$ for all $\bar{\bar{x}} \subset \bar{\bar{X}}$. It follows that $\bar{U}_1(\bar{y}) \in \bar{U}_1(\Gamma)_{LC} = \bar{U}_1(\Gamma)_C$, which concludes the proof.

2.7 COROLLARY. *If U is $\aleph_a$-compact and if there exists a set of power $\leq \aleph_a$ total on $U(X)_C$, then $\bar{\bar{U}}(\bar{\bar{X}}) \subset U(X)_C$ which is $\aleph_a$-separable.*

Given $\bar{\bar{x}} \in \bar{\bar{X}}$, let π be any finite subset of $\bar{Y}$. Then $[\pi]$ is a directed set ($\leq$ means set inclusion). By Helly's theorem (14), there exists for each π an $x_\pi \in X$ such that both $\bar{U}(\bar{y})x_\pi = \bar{\bar{x}}\bar{U}(\bar{y})$ for $\bar{y} \in \pi$ and $\| x_\pi \| \leq \| \bar{\bar{x}} \| + 1$.

[2] A subset of X will be called $\aleph_a$-separable if it contains a strongly dense set of power $\leq \aleph_a$.

As $U(X_1)$ is $\aleph$-compact for all $\aleph$ (Lemma 2.5), there exists a $y_0 \epsilon U(X)_C$ (Lemma 1.3) for which $\bar{\bar{x}}\bar{U}(\bar{y}) = \lim_{\pi} \bar{y}U(x_\pi) = \bar{y}(y_0)$ for all $\bar{y} \epsilon \bar{Y}$. By definition of $\bar{\bar{U}}$, $\bar{\bar{U}}(\bar{\bar{x}}) = y_0$. $U(X)_C$ is $\aleph_a$-separable, by Theorem 2.6, since $\bar{U}$ satisfies the same conditions that U satisfies.

2.8 THEOREM. *If $\{U_n\}$ is a sequence of $\aleph_a$-compact transformations and $\lim_n || U_n - U_0 || = 0$, then U_0 is $\aleph_a$-compact.*

Suppose $[x_\pi \mid \pi \epsilon$ directed set of power $\leq \aleph_a] \subset X_1$. Let $X_0 \equiv [x_\pi]_{LC}$. Then there will exist a subset of $\bar{Y}$ of power $\leq \aleph_a$ total on $U(X_0)_C$ (see the argument used in Theorem 1.9). By Lemma 2.5 we infer that $U_n((X_0)_1)$ is $\aleph$-closed. Let $[\alpha]$ be the class of all pairs $(\bar{y}, n)$ where $\bar{y} \epsilon \bar{Y}$ and n is one of the integers $(0, 1, 2, \cdots)$. We can now apply Lemma 1.8 to the numbers $[a_{\alpha\pi} = \bar{y}U_n(x_\pi)]$. Hence there exists a directed set $[\sigma]$ such that $\lim_{\sigma} \bar{y}U_n(x_{\pi_\sigma})$ exists and

$$\underline{\lim_{\pi}}\, \bar{y}U_n(x_\pi) \leq \lim_{\sigma} \bar{y}U_n(x_{\pi_\sigma}) \leq \overline{\lim_{\pi}}\, \bar{y}U_n(x_\pi)$$

for all $\bar{y} \epsilon \bar{Y}$ and $n = (0, 1, 2, \cdots)$. As $U_n((X_0)_1)^*$ is $\aleph$-compact for $n \geq 1$, there exists a $y_n \epsilon Y$ such that $\lim_{\sigma} \bar{y}U_n(x_{\pi_\sigma}) = \bar{y}(y_n)$ for all $\bar{y} \epsilon \bar{Y}$ and $n \geq 1$. Now given ε, there exists an N such that for $n, m \geq N$ $|| U_n - U_m || \leq \varepsilon$. Hence $| \bar{y}(y_n - y_m) | = \lim_{\sigma} | \bar{y}(U_n - U_m)x_{\pi_\sigma} | \leq || \bar{y} || \varepsilon$ so that $|| y_n - y_m || \leq \varepsilon$ for $n, m \geq N$. The sequence y_n will therefore converge to an element y_0 of Y. As $\lim_n \bar{y}U_n(x_\pi) = \bar{y}U(x_\pi)$ uniformly in π, the iterated limits theorem gives $\lim_{\sigma} \bar{y}\bar{U}(x_{\pi_\sigma}) = \bar{y}(y_0)$. By the above inequalities

$$\underline{\lim_{\pi}}\, \bar{y}U_0(x_\pi) \leq \bar{y}(y_0) \leq \overline{\lim_{\pi}}\, \bar{y}U_0(x_\pi)$$

for all $\bar{y} \epsilon \bar{Y}$. U_0 is therefore $\aleph_a$-compact.

3. The convex extension of $\aleph_0$ compact sets. The principal theorem of this section proves that the closed convex extension of an $\aleph_a$-compact set is likewise $\aleph_a$-compact. Making use of this theorem, we obtain a characterization of $\aleph_a$-compact subsets G of X which is independent of the separability of X. The set of numbers $[\bar{x}(x) \mid \bar{x} \epsilon \bar{X}_1, x \epsilon G]$ can be considered as a class of functions having the range $\bar{X}_1$ or as a class of functions having the range G. Our characterization of $\aleph_a$-compact sets exhibits these two function classes as duals of one another.

3.1 THEOREM. *If G is an $\aleph_0$-compact subset of a separable Banach space X, then $U(\bar{x}) = [\bar{x}(x) \mid x \epsilon G]$ on $\bar{X}$ to M_G is $\aleph_0$-compact.*

Since G is bounded (by Corollary 1.7), $[\bar{x}(x) \mid x \in G]$ belongs to M_G. Suppose that $\{\bar{x}_n\}$ is a sequence of functionals bounded by one. As X is separable, $\bar{X}_1$ is a compact metric space in its $\bar{w}(\bar{X})$ topology (1, pp. 185-186). Hence there exists a subsequence, which we renumber for convenience, and an $\bar{x}_0 \in \bar{X}$ such that $\lim_n \bar{x}_n(x) = \bar{x}_0(x)$ for all $x \in X$. We will show that $\lim_n \bar{b}\bar{U}(\bar{x}_n) = \bar{b}\bar{U}(\bar{x}_0)$ or that $\lim_n \bar{b}\bar{U}\bar{y}_n = 0$ where $\bar{y}_n = \bar{x}_n - \bar{x}_0$ for all $\bar{b} \in \bar{M}_G$. By a theorem due to Hildebrandt (13) this is equivalent to $\lim_n \int_G \bar{y}_n(x)\,d\beta = 0$ for all finitely additive bounded set functions β on all subsets of G. Suppose this were not the case. Then by a theorem due to Banach (1, p. 219), there exists a sequence $\{x_i\} \subset G$ such that $\lim_n \overline{\lim_i} \mid \bar{y}_n(x_i) \mid > \varepsilon > 0$. Since G is $\aleph_0$-compact there exists a subsequence $\{x_{i_j}\}$ which converges weakly as elements to an $x_0 \in X$. Therefore $\overline{\lim_i} \mid \bar{y}(x_i) \mid \leq \mid \bar{y}(x_0) \mid$. However $\lim_n \bar{y}_n(x_0) = 0$, which is in contradiction to our supposition.

3.2 Corollary. *If G is an $\aleph_0$-compact subset of a Banach space X, then $[\Sigma_G a_x x \mid \Sigma_G \mid a_x \mid \leq 1]$ is $\aleph_0$-compact.*

Suppose $x_n = \sum_{i=1}^{\infty} a_i^n x_i^n$ where $\sum_{i=1}^{\infty} \mid a_i^n \mid \leq 1$ and $x_i^n \in G$. We shall show that $\{x_n\}$ is contained in an $\aleph_0$ compact subset of X. Let $X_0 = [x_i^n]_{LC}$ and $G_0 = G \cap X_0$. Then G_0 is $\aleph_0$-compact (by Lemma 1.3) and is contained in the separable space X_0. Define $T(a) = \Sigma_{G_0} a_x x$ where $\Sigma_{G_0} \mid a_x \mid < \infty$. As G is bounded (by Corollary 1.7), T is linear on l_{G_0} to X. $\bar{T}(\bar{x}) = [\bar{x}(x) \mid x \in G_0]$ on $\bar{X}$ to M_{G_0} is by Theorem 3.1 $\aleph_0$-compact. By Corollary 2.3, T is likewise $\aleph_0$-compact. Therefore $T[(l_{G_0})_1]$ which contains $\{x_n\}$ is $\aleph_0$-compact.

Krein (15) has demonstrated the slightly weaker result that the convex extension of an $\aleph_0$ compact subset of a separable Banach space is likewise $\aleph_0$-compact. His proof made use of the Fubini Theorem. The following lemma is due to Smulian (25). For the sake of completeness we shall give another proof of this result.

3.3 Lemma. *If G is $\aleph_0$-compact, then G_C is $\aleph_0$-compact.*

If $\{x_n\} \subset G_C$, then there exist y_n in G such that $\| x_n - y_n \| < 1/n$. As G is $\aleph_0$-compact, there exists an x_0 for which

$$\underline{\lim_n}\, \bar{x}(y_n) \leq \bar{x}(x_0) \leq \overline{\lim_n}\, \bar{x}(y_n)$$

for all $\bar{x} \in \bar{X}$. Now if $\| \bar{x} \| \leq 1$ and $n \geq k$, then $\bar{x}(x_n) - 1/k \leq \bar{x}(y_n) \leq \bar{x}(x_n) + 1/k$. Thus

$$\underline{\lim_n}\, \bar{x}(x_n) - 1/k \leq \underline{\lim_n}\, \bar{x}(y_n) \leq \bar{x}(x_0) \leq \overline{\lim_n}\, \bar{x}(y_n) \leq \overline{\lim_n}\, \bar{x}(x_n) + 1/k$$

for all k. In other words

$$\underline{\lim_n}\, \bar{x}(x_n) \leq \bar{x}(x_0) \leq \overline{\lim_n}\, \bar{x}(x_n).$$

Hence by Theorem 1. 6, G_C is $\aleph_0$-compact.

3. 4 THEOREM. *If G is $\aleph_a$-compact, then G_C is $\aleph_a$-compact.*

If $[x_\pi \mid \pi \in$ directed set of power $\leq \aleph_a] \subset G_C$, then there exist elements $x_\pi{}^n$ in G such that $\| x_\pi - x_\pi{}^n \| < 1/n$. As G is $\aleph_a$-compact, there exist x^n in X for which

$$\underline{\lim_\pi}\, \bar{x}(x_\pi{}^n) \leq \bar{x}(x^n) \leq \overline{\lim_\pi}\, \bar{x}(x_\pi{}^n)$$

for all $\bar{x} \in \bar{X}$. It follows that

$$\underline{\lim_\pi}\, \bar{x}(x_\pi) - 1/n \leq \bar{x}(x^n) \leq \overline{\lim_\pi}\, \bar{x}(x_\pi) + 1/n$$

for all $\bar{x} \in \bar{X}_1$. By Lemma 1. 3 x_n belongs to the closed convex extension of G which, according to Corollary 3. 2 and Lemma 3. 3, is $\aleph_0$-compact. Hence we can find an $x_0 \in X$ such that

$$\underline{\lim_n}\, \bar{x}(x^n) \leq \bar{x}(x_0) \leq \overline{\lim_n}\, \bar{x}(x^n)$$

for all $\bar{x} \in \bar{X}$. Finally

$$\underline{\lim_\pi}\, \bar{x}(x_\pi) \leq \underline{\lim_n}\, \bar{x}(x^n) + 1/n \leq \bar{x}(x_0) \leq \overline{\lim_n}\, \bar{x}(x^n) - 1/n \leq \overline{\lim_\pi}\, \bar{x}(x_\pi)$$

for all $\bar{x} \in \bar{X}_1$.

We remark that the same argument shows G $\aleph_a$-closed to imply G_C $\aleph_a$-closed. The only lemmas used in the proof concerned $\aleph_0$-compactness which is equivalent to $\aleph_0$-closedness.

3. 5 LEMMA. *If G is $\aleph$-compact, then $[\Sigma_G a_x x \mid \Sigma_G \mid a_x \mid \leq 1]$ is $\aleph$-compact.*

Let $S_n \equiv [\sum_{i=1}^{n} c_i x_i \mid \sum^{n} \mid c_i \mid \leq 1, x_i \in G]$. We first show that S_n is $\aleph$-compact. Suppose $[x_\pi \mid \pi \in$ directed set$] \subset S_n$. That is $x_\pi = \sum_{i=1}^{n} c_i{}^\pi x_i{}^\pi$ where $\sum_{i=1}^{n} \mid c_i{}^\pi \mid \leq 1$ and $x_i{}^\pi \in G$. Let $a_{\bar{x}\pi}$ be the point in a $2n$-dimensional euclidean space having the coördinates $(\bar{x}(x_1{}^\pi), \bar{x}(x_2{}^\pi), \cdots, \bar{x}(x_n{}^\pi), c_1{}^\pi, c_2{}^\pi, \cdots, c_n{}^\pi)$. The set $[a_{\bar{x}\pi} \mid \pi \in [\pi]]$ is contained in a bounded and hence bicompact subset of this space. We can therefore apply Lemma 1. 8 and obtain a directed set $[\sigma]$

such that $\lim_{\sigma} \bar{x}(x_i^{\pi_\sigma})$ exists for each i and $\bar{x} \in \bar{X}$, and $\lim_{\sigma} c_i^{\pi_\sigma} = c_i$ exists for each i. As G is $\aleph$-compact, there exist $x_i \in X$ such that $\lim_{\sigma} \bar{x}(x_i^{\pi_\sigma}) = \bar{x}(\bar{x}_i)$ for all $\bar{x} \in \bar{X}$. Consequently $\lim_{\sigma} \bar{x}(c_i^{\pi_\sigma} x_i^{\pi_\sigma}) = \bar{x}(c_i x_i)$ and hence

$$\lim_{\sigma} \bar{x}(\sum_1^n c_i^{\pi_\sigma} x_i^{\pi_\sigma}) = \bar{x}(\sum_1^n c_i x_i).$$

Now for a given π_0 there exists a σ_0 such that $\pi_\sigma \geq \pi_0$ if $\sigma \geq \sigma_0$. Consequently if $x_0 = \sum_1^n c_i x_i$, then

$$\underline{\lim}_{\pi} \bar{x}(x_\pi) \leq \lim_{\sigma} \bar{x}(\sum_1^n c_i^{\pi_\sigma} x_i^{\pi_\sigma}) = \bar{x}(x_0) \leq \overline{\lim}_{\pi} \bar{x}(x_\pi).$$

We shall next show that $S = \Sigma S_n$ is $\aleph$-closed for all $\aleph$. Suppose therefore that $[x_\mu \mid \mu \in \text{ ordinal numbers } < \Omega] \subset S$. Let $\Pi_n = [\mu \mid x_\mu \in S_n]$. Then $\Sigma \Pi_n = [\mu \mid \mu < \Omega]$ and $\Pi_{n+1} \supset \Pi_n$. As S_n is $\aleph$-compact and Π_n is a directed set, there exists an $x_n \in S_C$ by Lemma 1.3 such that

$$\underline{\lim}_{\mu \in \Pi_n} \bar{x}(x_\mu) \leq \bar{x}(x_n) \leq \overline{\lim}_{\mu \in \Pi_n} \bar{x}(x_\mu)$$

for all $\bar{x} \in \bar{X}$. By Lemmas 3.5 and 3.3, S_C is $\aleph_0$-compact. Thus there exists an $x_0 \in X$ such that

$$\underline{\lim}_{n} \bar{x}(x_n) \leq \bar{x}(x_0) \leq \overline{\lim}_{n} \bar{x}(x_n).$$

For a given $\bar{x} \in \bar{X}$ and $\varepsilon > 0$, there exists a μ_0 such that $\mu \geq \mu_0$ implies $\bar{x}(x_\mu) \geq \underline{\lim}_{\mu} \bar{x}(x_\mu) - \varepsilon$. Further there exists an N such that $n \geq N$ implies $\mu_0 \in \Pi_n$ and hence $\underline{\lim}_{\mu \in \Pi_n} \bar{x}(x_\mu) \geq \underline{\lim}_{\mu} \bar{x}(x_\mu) - \varepsilon$. Therefore $\bar{x}(x_0) \geq \underline{\lim}_{\mu} \bar{x}(x_\mu) - \varepsilon$ for all ε. A symmetric argument gives

$$\underline{\lim}_{\mu} \bar{x}(x_\mu) \leq \bar{x}(x_0) \leq \overline{\lim}_{\mu} \bar{x}(x_\mu)$$

for all $\bar{x} \in \bar{X}$.

S is certainly convex and hence by Theorem 1.10, S is $\aleph$-compact for all $\aleph$. Finally the set $[\Sigma_G a_x x \mid \Sigma_G \mid a_x \mid \leq 1]$ is $\aleph$-compact by Theorem 3.4.

3.6 THEOREM. *If G is $\aleph_a$-compact, then $[\Sigma_G a_x x \mid \Sigma_G \mid a_x \mid \leq 1]$ is $\aleph_a$-compact.*

Suppose $[x_\pi \mid \pi \in \text{ directed set of power } \leq \aleph_a] \subset [\Sigma_G a_x x \mid \Sigma_G \mid a_x \mid \leq 1]$. Hence $x_\pi = \sum_{i=1}^{\infty} a_i^\pi x_i^\pi$ where $x_i^\pi \in G$. Let $X_0 \equiv [x_i^\pi]_{LC}$. By a familiar argument, there exists a set of power $\leq \aleph_a$ total on X_0. By Lemma 2.5, $G_0 \equiv G \cap X_0$

is $\aleph$-compact. Since $[\Sigma_{G_0} a_x x \mid \Sigma_{G_0} \mid a_x \mid \leq 1]$ is by Lemma 3. 5 $\aleph$-compact and contains $[x_\pi]$, there exists an $x_0 \epsilon X$ such that

$$\underline{\lim_{\pi}}\, \bar{x}(x_\pi) \leq \bar{x}(x_0) \leq \overline{\lim_{\pi}}\, \bar{x}(x_\pi)$$

for all $\bar{x} \epsilon \bar{X}$.

Combining this result with Theorem 3. 4, we obtain

3. 7 COROLLARY. *If G is $\aleph_a$-compact, then its closed convex extension is likewise $\aleph_a$-compact.*

We are now in a position to justify one of the introductory remarks, namely, G is $\aleph_a$-compact if and only if the $w(X)$ closure of G is $\aleph_a$-compact*. For if G is $\aleph_a$-compact, then its convex closed extension (which is weakly closed) is $\aleph_a$-compact and hence by Theorem 1. 9 this set is $\aleph_a$ compact*. The weak closure of G being a weakly closed subset of an $\aleph_a$-compact* set is likewise $\aleph_a$-compact*. The converse follows from Theorem 1. 9.

If G is a bounded subset of X and β a finitely additive set function on all subsets of G, we shall define $\int_G x\, d\beta$ to be an element x_0 of X for which $\bar{x}(x_0) = \int_G \bar{x}(x)\, d\beta$ for all $\bar{x} \epsilon \bar{X}$.

3. 8 THEOREM. *The following are equivalent:*

1). *G is $\aleph$-compact.*

2). *$(cv(G))_C$ is $\aleph$-compact.*

3). *$\int_G x\, d\beta$ exists and is unique for all β.*

4). *$\lim_{\sigma} \bar{x}_\sigma(x) = 0$ for all $x \epsilon X$ implies $\int_G \bar{x}_\sigma(x) d\beta \to 0$ for all β; and G is bounded.*

5). *$U(\bar{x}) = [\bar{x}(x) \mid x \epsilon G]$ is $\aleph$-compact on $\bar{X}$ to M_G for all $\aleph$.*

We shall show that 1) $\to$ 2) $\to$ 3) $\to$ 4) $\to$ 5) $\to$ 1).

That 1) implies 2) is a restatement of Corollary 3. 7. By Lemma 3. 5, $S \equiv [\sum_G a_x x \mid \sum_G \mid a_x \mid \leq 1]$ is $\aleph$-compact. Let π be a subdivision of G into a finite number of subsets $\{\tau_i\}$. Define $x_\pi = \Sigma_\pi x_i \beta(\tau_i)$ where x_i is some element of τ_i. Then $\int_G \bar{x}(x) d\beta = \lim_{\pi} \bar{x}(x_\pi)$. As $[x_\pi] \subset S$, it follows that there exists an $x_0 \epsilon X$ such that

$$\bar{x}(x_0) = \lim \bar{x}(x_\pi) = \int_G \bar{x}(x) d\beta$$

for all $\bar{x} \epsilon \bar{X}$. If $\bar{x}(x'_0) = \int_G \bar{x}(x) d\beta$ for all $\bar{x} \epsilon \bar{X}$, then $\bar{x}(x_0 - x'_0) = 0$ for

all $\bar{x} \in \bar{X}$ which implies $x_0 = x'_0$. Therefore 2) implies 3). That 3) implies 4) is obvious.

If G is bounded, we can define the linear transformation $T(a) = \sum_G a_x x$ where $\sum_G |a_x| < \infty$ on l_G to X. Then $\bar{T}(\bar{x}) = [\bar{x}(x) | x \in G]$ on $\bar{X}$ to M_G. We shall show that $\bar{T}$ is $\aleph$-compact as a consequence of 4). Given $[\bar{x}_\pi | \pi \in$ directed set$] \subset \bar{X}_1$. Applying Lemma 1.8 to the numbers $[\bar{x}_\pi(x) | x \in X, \pi \in [\pi]]$ we obtain a directed set $[\sigma]$ such that $\lim_\sigma \bar{x}_{\pi\sigma}(x)$ exists, and given π_0 there exists a σ_0 such that $\pi\sigma \geq \pi_0$ for all $\sigma \geq \sigma_0$. Now $\lim_\sigma \bar{x}_{\pi\sigma}(x)$ defines an element $\bar{x}_0$ of $\bar{X}$, namely, $\bar{x}_0(x) = \lim_\sigma \bar{x}_{\pi\sigma}(x)$. For every $\bar{b} \in \bar{M}_G$, $\bar{b}\bar{T}(\bar{x}) = \int_G \bar{x}(x)\, d\beta$. By 4) $\lim_\sigma \bar{b}\bar{T}(\bar{x}_{\pi\sigma}) = \bar{b}\bar{T}(\bar{x}_0)$ and hence $\lim_\pi \bar{b}\bar{T}(\bar{x}_\pi) \leq \bar{b}\bar{T}(\bar{x}_0) \leq \overline{\lim_\pi}\, \bar{b}\bar{T}(\bar{x}_\pi)$ which proves 5). By Corollary 2.3, T is likewise $\aleph$-compact for all $\aleph$. Therefore $G \subset T((l_G)_1)$ is $\aleph$-closed. This concludes the proof.

An application of this theorem gives a criterion for the interchange of order of integration for finitely additive set functions. Let $f(s,t)$ on $S \times T$ have the property that $[f(s,.) | s \in S]$ is an $\aleph$-compact subset of M_T. Then for any finitely additive set function α on all subsets of S and any finitely additive set function β on all subsets of T,

$$\int_T \int_S f(s,t)\, d\alpha\, d\beta = \int_S \int_T f(s,t)\, d\beta\, d\alpha.$$

If G is an $\aleph_a$-compact subset of X and $H \subset \bar{X}$ is such that for $x \in G$

$$k_1 \|x\| \leq \sup_{\bar{x} \in H} |\bar{x}(x)| \leq k_2 \|x\| \qquad \text{where } k_1, k_2 > 0$$

then $V(x) = [\bar{x}(x) | \bar{x} \in H]$ maps G isomorphically into a subset $\mathfrak{G}$ of M_H. Clearly $\mathfrak{G}$ is likewise $\aleph_a$-compact (see Lemma 2.1). Further $U(\bar{x}) = [\bar{x}(x) | x \in G]$ maps H into a subset $\mathfrak{H}$ of M_G since G is bounded. The following theorem states that $\mathfrak{G} \equiv [.(x) | x \in \mathfrak{G}]$ is $\aleph_a$-compact if and only if $\mathfrak{H} \equiv [\bar{x}(.) | \bar{x} \in H]$ is $\aleph_a$-compact. In this sense the sets $\mathfrak{G}$ and $\mathfrak{H}$ are the duals of one another. To make this duality more striking, let us consider the real valued bounded function $f(s,t)$ on $S \times T$. Then $\mathfrak{G} \equiv [f(.,t) | t \in T]$ is a subset of M_S and $\mathfrak{H} \equiv [f(s,.) | s \in S]$ is a subset of M_T. Using these definitions of $\mathfrak{G}$ and $\mathfrak{H}$ we have

3.9 THEOREM. *$\mathfrak{G}$ is $\aleph_a$-compact if and only if $\mathfrak{H}$ is $\aleph_a$-compact.*

By symmetry we need only show that if $\mathfrak{G}$ is $\aleph_a$-compact, then $\mathfrak{H}$ is $\aleph_a$-compact. Now $W(a) = \sum_T a_t f(.,t)$ where $\sum_T |a_t| < \infty$ is a linear trans-

formation on l_T to M_S since $\mathfrak{G}$ is bounded. As $\mathfrak{G}$ is $\aleph_a$-compact, Theorem 3.6 implies that $W((l_T)_1)$ is $\aleph_a$-compact. Hence W is $\aleph_a$-compact. By Theorem 2.2, $\bar{W}$ on $\bar{M}_S$ to M_T is $\aleph_a$-compact. Since $\bar{x}_s(x) = x(s)$ on M_S to real numbers is an element of $(\bar{M}_S)_1$, it follows that $\mathfrak{H} \equiv [\bar{W}(\bar{x}_s) = f(s, \cdot) \mid s \in S]$ is $\aleph_a$-compact which was to be proved.

3.10 COROLLARY. *A necessary and sufficient condition that G be $\aleph_a$-compact is that both* 1). *G be bounded, and* 2). *$U(\bar{x})$ on a determining manifold*[3] *Γ in $\bar{X}$ to M_G be $\aleph_a$-compact.*

In the above notation let H be the unit sphere in Γ. Then $\mathfrak{H} = U(H)$ is by Theorem 3.9 $\aleph_a$-compact. U is therefore $\aleph_a$-compact on Γ to M_G. Conversely if U is $\aleph_a$-compact, then $\mathfrak{H}$ is $\aleph_a$-compact. By Theorem 3.9, $\mathfrak{G}$ is then $\aleph_a$-compact and consequently $G = V^{-1}(\mathfrak{G})$ is $\aleph_a$-compact.[4]

We conclude this section with the following theorem:

3.11 THEOREM. *If G is $\aleph_a$-compact and there exists a subset of $\bar{X}$ of power $\leq \aleph_a$ total on G_{LC}, then G_{LC} is $\aleph_a$-separable.*

Let $T(a) = \sum_G a_x x$ where $\sum_G |a_x| < \infty$. T is linear limited on l_G to X and by Corollary 3.10 T is $\aleph_a$-compact. Hence by Theorem 2.2, T is $\aleph_a$-compact. As $G_{LC} = (T(l_G))_C$, the conclusion follows immediately from Corollary 2.7.

4. $\aleph$-compact sets. In this section we shall consider certain restrictions on the geometry of the space under which $\aleph_a$-compactness implies $\aleph$-compactness. Lemma 2.5 is an instance of this type of theorem. The principal problem here is still open. This is the problem of determining whether or not $\aleph_0$ compactness always implies $\aleph$-compactness. A subproblem to this is that of determining whether or not every space whose unit sphere is $\aleph_0$ compact is necessarily reflexive.

4.1 *Definition.* A *face* is a convex set of constant norm.

Milman (17) has introduced considerations of the type employed in the following theorem. In view of Theorem 3.11, this theorem is an extension of Lemma 2.5.

4.2 THEOREM. *If G is $\aleph_a$-compact and if there exist strongly dense subsets in every face of power $\leq \aleph_a$, then G is $\aleph$-compact.*

By Corollary 3.7, $(cv(G))_C$ is $\aleph_a$-compact. Theorem 1.10 reduces the

---

[3] A determining manifold Γ in $\bar{X}$ is a closed linear subset of $\bar{X}$ such that l. u. b. $[|\bar{x}(x)| \mid \bar{x} \in \Gamma, \ ||\bar{x}|| \leq 1] = ||x||$ for all $x \in X$.

[4] In this case V is an isometric transformation on X to a closed linear subset of M_H on which V^{-1} is defined.

problem to that of proving a closed convex set G which is $\aleph_a$-compact to be $\aleph$-closed for all $\aleph$. Suppose $[x_\mu \mid \mu \epsilon$ ordinals $< \Omega] \subset G$. Define $F_\mu \equiv cv[x_\lambda \mid \lambda \geq \mu]$ and $A_\mu = (F_\mu)c$. We first show that the A_μ have a common point. Since $A_\mu \subset G$ is $\aleph_0$-compact, convex, and weakly closed, it contains an element y_μ such that $d_\mu = || y_\mu || = \text{g. l. b. } [|| x' || \mid x' \epsilon A_\mu]$. Now if $\mu_1 > \mu_2$ then $A_{\mu_1} \subset A_{\mu_2}$, and consequently $d_{\mu_1} > d_{\mu_2}$. Thus $\lim_\mu d_\mu = d$ exists. We have the two following cases: 1). There exists no μ_0 such that $d_\mu = d$ for all $\mu \geq \mu_0$. In this case we can obtain a spanning sequence μ_n such that $| d_{\mu_n} - d | < 1/n$. Then a subsequence of $\{y_{\mu_n}\}$ converges weakly to an x_0, which is easily shown to lie in $\bigcap A_\mu$. 2). There exists a μ_0 such that $d_\mu = d$ for all $\mu \geq \mu_0$. The set $[y_\mu \mid \mu \geq \mu_0]$ therefore lies in a face. By hypothesis there exists a strongly dense subset in this face of power $\leq \aleph_a$. Let X_0 be the linear closed extension of this face. By a familiar argument, there exists a set of power $\leq \aleph_a$ total on X_0. Therefore $G_0 = G \bigcap X_0$ is by Lemma 2. 5 $\aleph$-compact. Hence there exists an x_0 such that

$$\underline{\lim_\mu} \bar{x}(y_\mu) \leq \bar{x}(x_0) \leq \overline{\lim_\mu} \bar{x}(y_\mu)$$

for all $\bar{x} \epsilon \bar{X}$. If x_0 were not an element of A_λ, then we could find an $\bar{x} \epsilon \bar{X}$ and a constant c such that $\bar{x}(A_\lambda) \leq c$ and $\bar{x}(x_0) > c$ (8, p. 331). Thus $\bar{x}(y_\mu) \leq c$ for all $\mu \geq \lambda$, contrary to the above inequalities. In both cases we can therefore find an $x_0 \epsilon \bigcap A_\mu$. For a given ε and μ_0 there exist $\mu_1, \mu_2, \cdots, \mu_n \geq \mu_0$ such that $|| \Sigma c_i x_{\mu_i} - x_0 || < \varepsilon$ where $c_i \geq 0$ and $\Sigma c_i = 1$. For one of these μ_i's, $\bar{x}(x_\mu) \leq \bar{x}(x_0) + \varepsilon$ while for another $\bar{x}(x_\mu) \geq \bar{x}(x_0) - \varepsilon$. Hence

$$\underline{\lim_\mu} \bar{x}(x_\mu) \leq \bar{x}(x_0) \leq \overline{\lim_\mu} \bar{x}(x_\mu)$$

for all $\bar{x} \epsilon \bar{X}$.

A space is said to be *strictly convex* if every face contains but a single element. We obtain as a corollary

4. 3 Corollary. *If G is $\aleph_0$-compact and X is strictly convex, then G is $\aleph$-compact.*

Clarkson (7) has shown that every separable space is isomorphic to a strictly convex space.[5] By employing the same method, one can easily prove that any space on which there exists a total set of power $\leq \aleph_0$ is isomorphic to a strictly convex space. This, however, gives no more than Lemma 2. 5.

Let $\alpha(\tau)$ be a non-negative measure function on the sigma field $\mathfrak{T}$ of subsets of T. We assume that $T = \Sigma \tau_i$ where $\alpha(\tau_i) < \infty$. Define $\mathfrak{M}(\mathfrak{T})$ to be the class of all essentially bounded measurable functions on T. $\mathfrak{M}(\mathfrak{T})$ is a

[5] The author has shown that the space M_T (where the power of $T > \aleph_0$) is not isomorphic to a strictly convex space.—Added September 1942.

Banach space under the norm $||\phi|| = \text{ess. l. u. b. } [|\phi(t)| \, t \in T]$. Let us define

$$||\phi||_1 = \left[\frac{||\phi||^2}{2} + \sum_{i=1}^{\infty} \frac{\int_{\tau_i} |\phi(t)|^2 \, d\alpha}{2^{i+1} \cdot \alpha(\tau_i)} \right]^{1/2}$$

Then $||\phi||/2^{1/2} \leq ||\phi||_1 \leq ||\phi||$. Now

$$\frac{||\phi+\psi||^2}{2} + \Sigma \frac{\int_{\tau_i} |\phi+\psi|^2 \, d\alpha}{2^{i+1} \cdot \alpha(\tau_i)} \leq \frac{(||\phi|| + ||\psi||)^2}{2} + \Sigma \frac{\int_{\tau_i} |\phi+\psi|^2 \, d\alpha}{2^{i+1} \cdot \alpha(\tau_i)}$$

$$= 2\,(||\phi||_1^2 + ||\psi||_1^2) - \left\{ \frac{(||\phi|| - ||\psi||)^2}{2} + \Sigma \frac{\int_{\tau_i} |\phi-\psi|^2 \, d\alpha}{2^{i+1} \cdot \alpha(\tau_i)} \right\}.$$

Thus if $||\phi+\psi||_1 = 2||\phi||_1 = 2||\psi||_1$, then $\int_{\tau_i} |\phi-\psi|^2 \, d\alpha = 0$ for all i. That is $\phi(t) = \psi(t)$ a. e. The renormed $\mathfrak{M}(\mathfrak{T})$ is therefore strictly convex and isomorphic to the original $\mathfrak{M}(\mathfrak{T})$. Making use of Corollary 4. 3 we have

4. 4 THEOREM. *If G is an $\aleph_0$-compact subset of $\mathfrak{M}(\mathfrak{T})$, then G is $\aleph$-compact.*

4. 5 THEOREM. *If G is an $\aleph_0$-compact subset of $\mathfrak{L}(\mathfrak{T})$, then G is $\aleph$-compact.*[6]

Dunford and Pettis (8, p. 376) have shown that $G \subset \mathfrak{L}(\mathfrak{T})$ is $\aleph_0$-compact if and only if

1). G is bounded,

2). given $\varepsilon > 0$ there is a $\delta > 0$ such that $\alpha(\tau) < \delta$ implies $|\int_{\tau} \phi \, d\alpha| < \varepsilon$ for every $\phi \in G$,

3). for each decomposition $\{\tau_i\}$ of T $\lim_{m\to\infty} \sum_{m}^{\infty} |\int_{\tau_i} \phi \, d\alpha| = 0$ uniformly over G.

We shall now show that these conditions are sufficient to imply that G is $\aleph$-compact. Suppose $[\phi_\pi \mid \pi \in \text{directed set}] \subset G$. Let $a_{\tau\pi} = \int_{\tau} \phi_\pi \, d\alpha$. Since G is bounded, we can apply Lemma 1. 8 and obtain a directed set $[\sigma]$ for which $\lim_{\sigma} a_{\tau\pi\sigma} = a_\tau$ exists for all $\tau \in \mathfrak{T}$. a_τ is clearly finitely additive on $\mathfrak{T}$

[6] $\mathfrak{L}(\mathfrak{T})$ is defined to be the class of all summable functions on T where $||\phi|| = \int_T |\phi(t)| \, d\alpha$ and $T = \Sigma\tau_i$ $[\alpha(\tau_i) < \infty]$.

and by condition 2) absolutely continuous with respect to α. If $\{\tau_n\}$ is a monotonic ascending sequence of sets $\tau_n \in \mathfrak{T}$ and $\lim \tau_n = \tau$, then by condition 3) $\lim_n a_{\tau_n \pi} = a_{\tau\pi}$ uniformly in π. Hence by the familiar iterated limits theorem $\lim_n a_{\tau_n} = a_\tau$. a_τ is therefore completely additive and by the Radon-Nikodym theorem the indefinite integral of an element ϕ of $\mathfrak{L}(\mathfrak{T})$. If ψ is a bounded step function on $\mathfrak{T}$, then $\lim_\sigma \int_T \psi\phi_{\pi\sigma}\, d\alpha = \int_T \psi\phi\, d\alpha$. Since the bounded step functions are strongly dense in $\mathfrak{M}(\mathfrak{T}) = \overline{\mathfrak{L}(\mathfrak{T})}$, it follows that $\lim_\sigma \int_T \psi\phi_{\pi\sigma}\, d\alpha = \int_T \psi\phi\, d\alpha$ for all $\psi \in \mathfrak{M}(\mathfrak{T})$. Finally, for every π_0, there exists a σ_0 such that $\pi\sigma \geq \pi_0$ for all $\sigma \geq \sigma_0$. Consequently

$$\underline{\lim}_\pi \int_T \psi\phi_\pi\, d\alpha \leq \int_T \psi\phi\, d\alpha \leq \overline{\lim}_\pi \int_T \psi\phi_\pi\, d\alpha$$

for all $\psi \in \mathfrak{M}(\mathfrak{T})$, which was to be proved.

For purposes of the next theorem, we define c_0 to be the class of all real valued functions $x(t)$ on an abstract class T which converge to zero for any sequence $\{t_i\} \subset T$. With $\| x \| = \text{l. u. b. } [| x(t)| \mid t \in T]$, c_0 is a Banach space.

4.6 THEOREM. *If G is an $\aleph_0$-compact subset of c_0, then G is $\aleph$-compact.*

We first show that if G is $\aleph_0$-compact then

1). G is bounded,

2). for all sequences $\{t_i\}$ of distinct elements of T $\lim_n \{\text{g. l. b. } [| x(t_i)| \mid i = 1, 2, \cdots, n]\} = 0$ uniformly over G.

1) follows from Corollary 1.7. Let us suppose that G $\aleph_0$-compact does not imply 2). Then there exists an $\varepsilon > 0$ and a sequence $\{t_i\}$ of distinct elements such that for every n we can find an $x_n \in G$ for which $| x_n(t_i)| \geq \varepsilon$, $i = 1, 2, \cdots, n$. As G is $\aleph_0$-compact, a subsequence $\{x_{n_i}\}$ converges weakly to some $x_0 \in c_0$. Since $x_{n_i}(t) \to x_0(t)$ for all $t \in T$, it follows that $| x_0(t_i)| \geq \varepsilon$ for all t_i. This is impossible for $x_0 \in c_0$.

We next show that conditions 1) and 2) are sufficient to make G $\aleph$-compact. Suppose $[x_\pi \mid \pi \in \text{directed set}] \subset G$. By 1), the numbers $x_\pi(t)$ are bounded. Applying Lemma 1.8, we obtain a directed set $[\sigma]$ such that $\lim_\sigma x_{\pi\sigma}(t) = a(t)$ exists for all $t \in T$. If $a(t)$ did not belong to c_0, there would exist an $\varepsilon > 0$ and a sequence $\{t_i\}$ of distinct points such that $| a(t_i)| \geq \varepsilon$ for all i. By 2), there exists an N such that g. l. b. $[| x(t_i)| \mid i = 1, 2, \cdots, N] < \varepsilon/2$, for all $x \in G$. This is contrary to the fact that we can find a σ_0 such that $\sigma \geq \sigma_0$

implies $| x_{\pi\sigma}(t_i) - a(t_i)| < \varepsilon/2$ for $i = 1, 2, \cdots, N$. Hence $a(t) \in c_0$. Since $\bar{c}_0 \equiv l_T$, it is easy to show that $\lim_\sigma \bar{x}(\pi_\sigma) = \bar{x}(a)$ for all $\bar{x} \in l_T$. Finally, for every π_0, there exists a σ_0 such that $\pi_\sigma \geq \pi_0$ for all $\sigma \geq \sigma_0$. Consequently

$$\underline{\lim}_\pi \bar{x}(x_\pi) \leq \bar{x}(a) \leq \overline{\lim}_\pi \bar{x}(x_\pi)$$

for all $\bar{x} \in l_T$.

4.7 THEOREM. *If G is an $\aleph_0$-compact subset of l_T, then G is $\aleph$-compact.*

To prove this, we show that l_T is isomorphic to a strictly convex space. The result then follows by Corollary 4.3. Define

$$|| a ||_1 = [(\sum_T | a_t |)^2 + \sum_T | a_t |^2]^{1/2}.$$

Then $|| a || \leq || a ||_1 \leq 2^{1/2} \cdot || a ||$. The remainder of the proof is similar to that of Theorem 4.4.

The corresponding theorem for $M_T \equiv l_T$ is, to our knowledge, unproved. Such a theorem would settle the general problem since any Banach space is equivalent to a subspace of an M_T space. We shall conclude this section with a characterization of $\aleph_0$-compact subsets of M_T in the hope that this characterization may be found useful in attacking the general problem.

4.8 THEOREM. *$G \subset M_T$ is $\aleph_0$-compact if and only if G is bounded and for all sequences $\{t_i\}$ there exists a subsequence $\{t_{i_j}\}$ and an $\bar{x} \in \bar{M}_T$ such that for no $\varepsilon > 0$ and no subsequence $\{t_{i_{j_k}}\}$ is*

$$\bigcap_{k=1}^n [x \mid | x(t_{i_{j_k}}) - \bar{x}(x)| > \varepsilon] \neq 0$$

for all n.

We can apply Theorem 3.9 to $f(t, x) \equiv x(t)$ on $T \times G$. The following statements are equivalent: 1). $G \equiv \mathfrak{G} \equiv [f(., x) \mid x \in G]$ is $\aleph_0$-closed; 2). $\mathfrak{H} \equiv [f(t, .) \mid t \in T] \subset M_G$ is $\aleph_0$-compact; 3). Given any sequence $\{t_i\}$ there exists a subsequence $\{t_{i_j}\}$ and an element $y(x)$ of M_G such that $\lim_j \overline{\lim_n} \mid x_n(t_{i_j}) - y(x_n)| = 0$ for all sequences $\{x_n\} \subset G$. The last equivalence follows from a theorem due to Banach (1, p. 219). Now $y(x)$ can be extended over M_T to be the functional $\bar{x}(x) = \operatorname{Lim}_j x(t_{i_j})$ where Lim stands for the Banach generalized limit function (1, p. 34). 3) can be restated as: given any sequence $\{t_i\}$ there exists a subsequence $\{t_{i_j}\}$ and an $\bar{x} \in \bar{M}_T$ such that for no $\varepsilon > 0$ and no subsequence $\{t_{i_{j_k}}\}$ of $\{t_{i_j}\}$ is $\underline{\lim}_n \mid x_n(t_{i_{j_k}}) - \bar{x}(x_n)| > \varepsilon$ for some sequence $\{x_n\} \subset G$. The required condition clearly implies this. Suppose the converse were not true. Then there would exist an $\varepsilon > 0$ and a subsequence $\{t_{i_{j_k}}\}$ of $\{t_{i_j}\}$ such that $\bigcap_{k=1}^n [x \mid |x(t_{i_{j_k}}) - \bar{x}(x)| > \varepsilon] \neq 0$ for all n. Choose

x_n out of $\bigcap_{k=1}^{n}[x \mid |x(t_{i_{j_k}}) - \bar{x}(x)| > \varepsilon]$. Then $\overline{\lim_n} |x_n(t_{i_{j_k}}) - \bar{x}(x_n)| \geq \varepsilon$, in contradiction to our assumption.

5. Integral representations of absolutely continuous set functions. The principal theorems of this section deal with an extension of the Radon-Nikodym theorem from real valued set functions to certain abstract valued functions. Gelfand (11), Clarkson (7), Pettis (19), and Dunford and Pettis (8) have obtained many results on this problem. By assuming that the set function lies in an "essentially" $\aleph_0$-compact subset, we have been able to extend these results to sigma fields which are not separable. As an application, we obtain the general form of the weakly completely continuous transformation on $\mathfrak{L}(\mathfrak{T})$ to an arbitrary Banach space and prove that it is separable valued.

We shall consider an abstract class T of elements t possessing a sigma field $\mathfrak{T}$ of subsets τ of T. $\alpha(\tau)$ will be a single-valued, non-negative, completely additive measure function of $\mathfrak{T}$. We shall assume that there exists a denumerable decomposition $\{\tau_i\}$ of T such that $\alpha(\tau_i) < \infty$. It will be convenient to designate by $|\tau|$ the value $\alpha(\tau)$. We shall define the class of additive X valued set functions

$$V^\infty(X) \equiv [x(\tau) \mid ||x(\tau)|| \leq M \cdot |\tau|, \tau \in \mathfrak{T}, |\tau| < \infty].$$

Finally we shall define $B^\infty(X)$ to be the class of weakly measurable point functions $y(t)$ on T to X whose values are contained in a *separable* $\aleph$-compact subspace of X. With the norm $||x(.)|| = \text{l. u. b.}\,[(||x(\tau)||/|\tau|) \mid \tau \in \mathfrak{T}, 0 < |\tau| < \infty]$ for $V^\infty(X)$ and $||y(.)|| = \text{ess. l. u. b.}\,[||y(t)|| \mid t \in T]$ for $B^\infty(X)$, both of these classes become Banach spaces.

Integration with respect to α will be realized by means of the Bochner integral (2). Since $y(t) \in B^\infty(X)$ is contained in a separable subspace of X, $y(t)$ is measurable (20, theorem 1.1) and hence integrable on all sets $\tau \in \mathfrak{T}$ of finite measure.

5.1 THEOREM. *If $x(.) \in V^\infty(X)$ and $S \equiv [(x(\tau)/|\tau| \mid 0 < |\tau| < \infty]$ is $\aleph_0$-compact, then there exists a $y(.) \in B^\infty(X)$ such that $x(\tau) = \int_\tau y(t)\, d\alpha$ for all τ of finite measure.*

If R is the space of real numbers, then by the Radon-Nikodym theorem there exists for each $f(.) \in V^\infty(R)$ a $g(.) \in B^\infty(R)$ such that $f(\tau) = \int_\tau g(t)\, d\alpha$ for all τ of finite measures. $V(f(.)) = g(.)$ defines an equivalence between $V^\infty(R)$ and $B^\infty(R)$. Hence by Theorem 4.4, $\aleph_0$ compact subsets of $V^\infty(R)$ are $\aleph$-compact. Define $W(\bar{x}) = [(\bar{x}(x(\tau))/|\tau|) \mid \tau \in \mathfrak{T}, 0 < |\tau| < \infty]$ on $\bar{X}$ to $V^\infty(R)$. Since $S \equiv [(x(\tau)/|\tau|) \mid \tau \in \mathfrak{T}, 0 < |\tau|$

$< \infty]$ is $\aleph_0$-compact, it follows by Corollary 3.10 that W is $\aleph_0$-compact. As $W(\bar{X}_1)$ is an $\aleph_0$-compact subset of $V^\infty(R)$, it must be $\aleph$-compact. Hence by this same corollary S is $\aleph$-compact.

Let π be a finite number of disjoint sets τ of $\mathfrak{T}$ for which $0 < |\tau| < \infty$. $\pi_1 \geq \pi_2$ will mean that $\Sigma\tau^1 \supset \Sigma\tau^2$ and that every set $\tau^1 \in \pi_1$, is either a subset of some $\tau^2 \in \pi_2$ or that τ^1 is disjoint from every $\tau^2 \in \pi_2$. For each π, we define $y_\pi(t) = x(\tau_i)/|\tau_i|$ for $t \in \tau_i$ and $y_\pi(t) = 0$ elsewhere. Then if $V(\bar{x}(x(.))) = g_x^-(t)$, $\lim_\pi |g_x^-(t) - \bar{x}(y_\pi(t))| = 0$ for *all* $t \in T$ and uniformly for any set of finite measure. As $[y_\pi(t)] \subset S$ is $\aleph$-compact, there exists a function $y(t)$ such that $\bar{x}(y(t)) = \lim_\pi \bar{x}(y_\pi(t)) = g_x^-(t)$ for all $\bar{x} \in \bar{X}$. $y(t)$ is clearly weakly measurable. By Lemma 1.3, $y(t)$ belongs to $(cv(S))_C$ which is $\aleph$-compact by Theorem 3.6. Further, it is clear that $\bar{x}(x(\tau)) = \int_\tau g_x^-(t)\, d\alpha = \int_\tau \bar{x}(y(t))\, d\alpha$ for all τ of finite measure.

We shall now show that $y(t)$ is separable valued. Actually we shall prove a stronger result, namely if $|\tau_0| < \infty$, then $[x(\tau) | \tau \subset \tau_0]$ is compact valued. This is equivalent to the statement that $[\bar{x}(x(\tau)) | \tau \subset \tau_0, \bar{x} \in \bar{X}_1]$ is a compact subset of $M_{(\tau \subset \tau_0)}$ (21, theorem 3.1). As $(cv(S))_C$ is $\aleph$-compact, it follows from Corollary 3.10 that given any sequence $\{\bar{x}_n\} \subset \bar{X}_1$ there exists a subsequence $\{\bar{x}_{n'}\}$ and an $\bar{x}_0 \in \bar{X}$ such that $\lim_n \bar{x}_{n'}(x) = \bar{x}_0(x)$ for all $x \in (cv(S))_C$. $y(t)$ is bounded. Consequently $\bar{x}_{n'}(x(\tau)) = \int_\tau \bar{x}_{n'}(y(t))\, d\alpha$ converges to $\bar{x}_0(x(\tau))$ uniformly for all $\tau \subset \tau_0$. Now $T = \Sigma\tau_i$ where $|\tau_i| < \infty$. Consequently $[x(\tau) | \tau \in \mathfrak{T}, |\tau| < \infty]$ is separable valued. But S and hence $(cv(S))_C$ is contained in the closed linear extension of this set which is clearly a separable subspace of X. Therefore $y(t)$ is separable valued. Thus $y(.) \in B^\infty(X)$. That $x(\tau) = \int_\tau y(t)\, d\alpha$ follows from $\bar{x}(x(\tau)) = \int_\tau \bar{x}(y(t))\, d\alpha$ for all $\bar{x} \in \bar{X}$.

5.2 THEOREM. *The general form of the weakly completely continuous transformation U on $\mathfrak{L}(\mathfrak{T})$ to X is*

$$U(\phi) = \int \phi(t) y(t)\, d\alpha \tag{1}$$

where $y(.) \in B^\infty(X)$ and $||y(.)|| = ||U||$. U is separable valued and $\aleph$-compact for all $\aleph$.

Let χ_τ be the characteristic function of the set $\tau \in \mathfrak{T}$. Define $x(\tau) = U(\chi_\tau)$ for all $\tau \in \mathfrak{T}$ such that $|\tau| < \infty$. It is clear that $x(.)$ belongs to $V^\infty(X)$ and that $[(x(\tau)/|\tau|) | \tau \in \mathfrak{T}, 0 < |\tau| < \infty]$ is $\aleph_0$-compact. Applying Theorem

5.1, we obtain $y(.) \in B^\infty(X)$ where $x(\tau) = \int_\tau y(t)d\alpha$ for all τ of finite measure. Now (1) is satisfied by all step functions in $\mathfrak{L}(\mathfrak{T})$. Since

$$2) \qquad \| \int \phi(t)y(t)d\alpha \| \leq \| \phi \| \cdot \| y(.) \|$$

and since all functions in $\mathfrak{L}(\mathfrak{T})$ can be approximated by step functions, (1) is valid in the general case. By (2) we have $\| U \| \leq \| y(.) \|$. As $(\| x(\tau) \| / | \tau | \leq \| U(\chi_\tau / | \tau |) \| \leq \| U \|$, it follows from the way in which $y(t)$ was found that $\| y(t) \| \leq$ l. u. b. $[(\| x(\tau) \| / | \tau |) \mid 0 < \tau < \infty] \leq \| U \|$. Hence $\| U \| = \| y(.) \|$. Further $\int \phi(t)y(t)d\alpha \in [y(t)]_{LC}$. Therefore U is separable valued. Conversely, it is easily seen that (1) defines a linear limited transformation. Suppose $\| \phi \| \leq 1$. Then $\int \phi(t)y(t)d\alpha$ is approximated by elements of the type $\Sigma\phi(t_i) \mid \tau_i \mid y(t_i)$ where $t_i \in \tau_i$ and $\{\tau_i\}$ is a decomposition of T. But these are contained in the set $[\sum_T a_t y(t) \mid \sum_T \mid a_t \mid \leq 1]$ which by Theorem 3.6 is $\aleph$-compact. Therefore $U((\mathfrak{L})_1)$ lies in the closure of this set which by Theorem 3.4 is also $\aleph$-compact. This concludes the proof.

An alternative proof for Theorems 5.1 and 5.2 will now be presented. This argument is not as direct, depending upon a long chain of results. Suppose $T = \Sigma\tau_i$ where τ_i are disjoint sets of $\mathfrak{T}$ of finite measure. Set

$$C_n \equiv [\phi(t) \mid \mid \phi(t) \mid \leq n \text{ for } t \in \sum_{i=1}^{n} \tau_i, \ \phi(t) = 0 \text{ elsewhere}].$$

As shown in Theorem 4.5, C_n is $\aleph_0$-compact. Suppose that U is $\aleph_0$-compact on $\mathfrak{L}(\mathfrak{T})$ to X. It can be shown that $U(C_n)$ is conditionally compact (21, theorem 5.5) and hence contained in a separable closed linear subspace X_n of X. Now ΣC_n is dense in $\mathfrak{L}(\mathfrak{T})$. Therefore $U(\mathfrak{L}) \subset (\Sigma X_n)_C$ which is a separable subspace of X. U is therefore separable valued. The general form of the weakly completely continuous *separable* transformation has been shown to be that given by Theorem 5.2 (21, theorem 5.4). From this one can easily obtain Theorem 5.1.

We remark that Theorem 5.2 can not be extended to the general space of Lebesgue integrable functions where T is not the sum of a denumerable number of sets of finite measure. For suppose $\mathfrak{L} = l_T$ where $\bar{T} \geq \aleph_1$. Let $X \equiv l_T{}^2 \equiv [x_t \mid \sum_T x_t{}^2 < \infty]$. Define $U(a) = x$ where $a_t = x_t$ for all $t \in T$. Now

$$\| U(a) \| = [\Sigma a_t{}^2]^{1/2} \leq [(\Sigma \mid a_t \mid)^2]^{1/2} = \| a \|.$$

U is therefore bounded and certainly not separable valued. Finally since l_T^2 is reflexive, U is $\mathfrak{N}$-compact.

5.3 LEMMA. *If $x(\tau)$ is a completely additive α-absolutely continuous set function on $\mathfrak{T}$ to X where $\alpha(T) < \infty$, then for each n there exists a decomposition $(\tau_0, \tau_1, \tau_2, \cdots)$ of T into disjoint sets of $\mathfrak{T}$ such that*

$$||x(\tau)|| \leq n|\tau| \quad \text{for all } \tau \subset \tau_0$$

and

$$||x(\tau)|| \geq n|\tau| \quad \text{for all } \tau \subset \tau_i \quad i = 1, 2, \cdots.$$

Given $\bar{x} \in \bar{X}$, it follows from a theorem due to Hahn (22, p. 32) that there exists a set $\tau_{\bar{x}} \in \mathfrak{T}$ such that $\bar{x}(x(\tau)) \geq n|\tau| \cdot ||\bar{x}||$ for all $\tau \subset \tau_{\bar{x}}$ and $\bar{x}(x(\tau)) \leq n \cdot |\tau| \cdot ||\bar{x}||$ for all $\tau \subset T - \tau_{\bar{x}}$. Let $O \equiv [\tau \mid \tau \subset \tau_{\bar{x}}$ for some $\bar{x}, |\tau| > 0]$. We define a transfinite sequence $(\tau_1, \tau_2, \cdots; \tau_\omega, \cdots, \tau_\lambda, \cdots)$ as follows: τ_1 is an arbitrary element of O. Suppose τ_λ to have been defined for all $\lambda < \mu$. If there exists τ such that $\tau \cdot \sum_{\lambda<\mu} \tau_\lambda = 0$ and $\tau \in O$, then set $\tau = \tau_\mu$. As α is bounded, $\alpha(\tau)$ cannot differ from zero on a nondenumerable number of disjoint sets. The sequence therefore contains only a denumerable set of elements. Let us put this set into one-to-one correspondence with the natural numbers. Define $\tau_0 = T - \sum_{i=1}^{\infty} \tau_i$. Then if $\tau \subset \tau_0$, it follows that $|\tau \cdot \tau_{\bar{x}}| = 0$ for all $\bar{x} \in \bar{X}$. Hence $\bar{x}(x(\tau)) \leq n|\tau| \cdot ||\bar{x}||$ for all $\bar{x} \in \bar{X}$, so that $||x(\tau)|| \leq n|\tau|$. On the other hand τ_i is a subset of some $\tau_{\bar{x}_i}$. Thus if $\tau \subset \tau_i$ then $\bar{x}_i(x(\tau)) \geq n|\tau| \cdot ||\bar{x}_i||$ and therefore $||x(\tau)|| \geq n|\tau|$.

Lemma 5.3 can obviously be extended to the case of $T = \Sigma\tau_i$ where τ_i is of finite measure. In the following work we will consider such T.

We define $x(\tau)$ to be of *p-bounded variation* for $p \geq 1$ if for all finite sets $\pi = (\tau_1, \tau_2, \cdots, \tau_n)$ of disjoint sets of finite measure different from zero

$$V_p(x(.)) = \text{l. u. b.}_{\pi} \left[\sum_{\pi} (||x(\tau_i)||^p / |\tau_i|^{p-1})\right] < \infty.$$

5.4 LEMMA. *If $x(\tau)$ is a completely additive α absolutely continuous set function on $\mathfrak{T}$ to X and of p-bounded variation $(p \geq 1)$, then there exist sets $\{\tau_n\}$ such that $\tau_n \supset \tau_{n+1}$, $\lim_n |\tau_n| = 0$, and $||x(\tau)|| \leq n|\tau|$ for all $\tau \subset T - \tau_n$.*

Applying Lemma 5.3 to n, we obtain the decomposition $(\sigma^0, \sigma^1, \sigma^2, \cdots)$ of T into disjoint sets such that $||x(\tau)|| \leq n|\tau|$ for all $\tau \subset \sigma^0$ and $||x(\tau)|| \geq n|\tau|$ for all $\tau \subset \sigma^i$ $(i = 1, 2, \cdots)$. As $x(\tau)$ is of p-bounded variation, there exists an M such that

$$M \geq \sum_{i=1}^{k} (||x(\sigma^i)||^p / |\sigma^i|^{p-1}) \geq n^p \cdot |\sum_{i=1}^{k} \sigma^i|.$$

Consequently if we define $\sigma_n = \sum_{i=1}^{\infty} \sigma^i$, then $\alpha(\sigma_n) \leq M/n^p$. Let $\sigma =$

$\sigma_n \cdot (T - \sigma_{n-1})$. Then $(n-1) \cdot |\sigma \cdot \sigma^i| \geq ||x(\sigma \cdot \sigma^i)|| \geq n \cdot |\sigma \cdot \sigma^i|$. Thus $|\sigma \cdot \sigma^i| = 0$ and hence $|\sigma| = 0$. Define $\tau_n = \sum_n^\infty \sigma_i$. Then $|\tau_n - \sigma_n| = 0$. The $\{\tau_n\}$ clearly satisfy the conditions of the theorem.

5.5 THEOREM. *If $x(\tau)$ is a completely additive α-absolutely continuous set function on T to X of p-bounded variation such that*

$$[(x(\tau)/|\tau|)|(||x(\tau)||)/|\tau| \leq n]$$

is $\aleph_0$-compact for all n, then there exists a measurable function $y(t)$ such that

$$x(\tau) = \int_\tau y(t)d\alpha \begin{cases} \text{for all } \tau \text{ if } p = 1 \\ \text{for all } \tau \text{ of finite measure if } p > 1 \end{cases}$$

and

$$V_p(x(.)) = \text{l. u. b.}_\pi [\sum_\pi (||x(\tau_i)||^p)/|\tau_i|^{p-1}] = \int_T ||y(t)||^p \, d\alpha.$$

We can apply Lemma 5.4 and obtain sets τ_n such that $\tau_n \supset \tau_{n+1}$, $\lim_n |\tau_n| = 0$, and $||x(\tau)|| \leq n|\tau|$ for all $\tau \subset T - \tau_n$. By Theorem 5.1 there exists for each n a $y_n(.) \subset B^\infty(X)$ such that $y_n(t) = 0$ for all $t \in \tau_n$ and $x_n(\tau) \equiv x(\tau - \tau \cdot \tau_n) = \int_\tau y_n(t)d\alpha$ for all τ of finite measure. Since the $y_n(t)$ are separable valued $y_n(t) = y_{n+1}(t)$ a. e. on $T - \tau_n$. We define $y(t) = y_n(t)$ for $t \in \tau_{n-1} - \tau_n$ where $T = \tau_0$. Then as $|\tau_n| \to 0$, $y(t)$ is measurable and $y(t) = y_n(t)$ a. e. on $T - \tau_n$.

Now if $z(t)$ is measurable and $\int_T ||z(t)||^p d\alpha < \infty$, one shows by step function approximations that $V_p(z(\tau)) = \int_T ||z(t)||^p d\alpha$ where $z(\tau) \int_\tau z(t)d\alpha$. Consequently $\int_T ||y_n(t)||^p d\alpha = V_p(x_n(.)) \leq V_p(x(.))$. Hence $\int_T ||y(t)||^p d\alpha = V_p(x(.))$. Since $||y_n(t) - y(t)||^p \leq ||y(t)||^p$, it follows that $\lim_n \int_T ||y_n(t) - y(t)||^p d\alpha = 0$. Therefore for all τ of finite measure $x(\tau) = \lim_n x_n(\tau) = \lim_n \int_\tau y_n(t)d\alpha = \int_\tau y(t)d\alpha$. If $p = 1$, this holds for all τ. For given a τ there exists a monotonic ascending sequence of sets τ_n of finite measure such that $\lim_n \tau_n = \tau$. $x(\tau) = \lim_n x(\tau_n) = \lim_n \int_{\tau_n} y(t)d\alpha = \int_\tau y(t)d\alpha$ when $p = 1$.

5.6 COROLLARY. *If $x(\tau)$ is a completely additive α-absolutely continuous*

set function on $\mathfrak{T}$ to a reflexive Banach space X of p-bounded variation, then there exists a measurable function $y(t)$ such that

$$x(\tau) = \int_\tau y(t)\,d\alpha \begin{cases} \text{for all } \tau \text{ if } p = 1 \\ \text{for all } \tau \text{ of finite measure if } p > 1 \end{cases}$$

and

$$V_p(x(.)) = \int_T || y(t) ||^p d\alpha.$$

We shall conclude this section with an application of Corollary 5.6 to the spaces $L^p(X)$ introduced by Bochner and Taylor (3). For this purpose we do not insist upon the usual decomposition of T into a denumerable sum of sets of finite measure. $L^p(X)$ is defined to be the class of all measurable functions $x(t)$ on T to X for which $|| x(.) || = (\int_T || x(t) ||^p dt)^{1/p}$ exists.

Let U be a linear functional on $L^p(X)$ where $p > 1$. Define $\chi(\tau, x) = x$ for all $t \in \tau$ and $\chi(\tau, x) = 0$ elsewhere. If τ is of finite measure then $\chi(\tau, x) \in L^p(X)$ and $U(\chi(\tau, x))$ exists. Following the reasoning of Bochner and Taylor (3, theorem 3.2), one can show that $U(\chi(\tau, x))$ is a linear functional $\bar{x}(\tau)$ on X, that $\bar{x}(\tau)$ is additive on sets of finite measure, that $|| U ||_q = V_q(\bar{x}(\tau))$ $(1/p + 1/q = 1)$, and that $U(x(.)) = \int d\bar{x}(\tau) \cdot x(t)$. Since $V_q(\bar{x}(\tau)) < \infty$, $\bar{x}(\tau)$ is completely additive and α-absolutely continuous on all subsets of a set of finite measure to $\bar{X}$. Further one can easily prove $V_q(\bar{x}(\tau)) < \infty$ implies the existence of a set τ_0 which has a denumerable decomposition into disjoint sets τ_i of finite measure and for which $\tau \cdot \tau_0 = 0$ gives $x(\tau) = 0$.

If we assume X to be reflexive, then $\bar{X}$ is likewise reflexive and we can apply Corollary 5.6 to the sets $\{\tau_i\}$. On each of these there will exist a function $\bar{x}_i(t)$ such that $\bar{x}(\tau \cdot \tau_i) = \int_\tau \bar{x}_i(t)\,dt$ and $V_q(\bar{x}(\tau \cdot \tau_i)) = \int_{\tau_i} || \bar{x}_i(t) ||^q dt$. Define $\bar{x}(t) = \bar{x}_i(t)$ for $t \in \tau_i$ $(i = 1, 2, \cdots)$ and $\bar{x}(t) = 0$ elsewhere. Then $\bar{x}(t)$ is measurable, $\bar{x}(\tau) = \int_\tau \bar{x}(t)\,d\alpha$ for all sets of finite measure, and $|| U ||^q = V_q(\bar{x}(.)) = \int_T || \bar{x}(t) ||^q d\alpha$. In other words $\overline{L^p(X)} = L^q(\bar{X})$.

5.7 THEOREM. *If X is reflexive, then $\overline{L^p(X)} = L^q(\bar{X})$ and hence $L^p(X)$ is reflexive.*

HARVARD UNIVERSITY.

REFERENCES

1. S. Banach, *Théorie des Opérations Linéaires*, Warsaw, 1932.
2. S. Bochner, "Integration von Funktionen deren Werte die Elemente eines Vektorraumes sind," *Fundamenta Mathematicae*, vol. 20 (1933), pp. 262-276.
3. S. Bochner and A. E. Taylor, "Linear functionals on certain spaces of abstractly-valued functions," *Annals of Mathematics*, 2, vol. 39 (1938), pp. 913-944.
4. N. Bourbaki, "Sur les espaces de Banach," *Comptes Rendus*, 206 (1938), pp. 1701-1704.
5. D. G. Bourgin, "Some properties of Banach spaces," *American Journal of Mathematics*, vol. 64 (1942), pp. 597-612.
6. Eduard Čech, "On bicompact spaces," *Annals of Mathematics*, vol. 38 (1939), pp. 823-846.
7. J. A. Clarkson, "Uniformly convex spaces," *Transactions of the American Mathematical Society*, vol. 40 (1936), pp. 396-414.
8. Nelson Dunford and B. J. Pettis, "Linear operations on summable functions," *Transactions of the American Mathematical Society*, vol. 47 (1940), pp. 323-392.
9. V. Gantmacher and V. Smulian, "Sur les espaces linéaires dont la sphère unitaire est faiblement compacte," *Comptes Rendus U. R. S. S.*, vol. 17 (1937), pp. 91-94.
10. V. Gantmacher, "Über schwache totalstetige Operatoren," *Recueil Mathématique*, vol. 7 (1940), pp. 301-308.
11. I. Gelfand, "Abstrakte Funktionen und lineare Operatoren," *Recueil Mathématique*, vol. 4 (1938), pp. 235-284.
12. H. Goldstine, "Weakly complete Banach spaces," *Duke Mathematical Journal*, vol. 4 (1938), pp. 125-131.
13. T. H. Hildebrandt, "On bounded linear functional operations," *Transactions of the American Mathematical Society*, vol. 36 (1934), pp. 868-875.
14. S. Kakutani, "Weak topology and regularity of Banach spaces," *Proceedings of the National Academy of Sciences of Tokyo*, vol. 25 (1939), pp. 169-173.
15. M. Krein, "Sur quelques questions de la géometrie des ensembles convexes situé dans un espace linéaire normé et complet," *Comptes Rendus U. R. S. S.*, vol. 14 (1937), pp. 5-7.
16. S. Mazur, "Über konvex Mengen in linearen normierten Räumen," *Studia Mathematica*, vol. 4 (1933), pp. 70-84.
17. D. Milman, "On some criteria for the regularity of spaces of the type (B)," *Comptes Rendus U. R. S. S.*, vol. 20 (1938), pp. 243-246.
18. E. H. Moore and H. L. Smith, "A general theory of limits," *American Journal of Mathematics*, vol. 44 (1922), pp. 102-121.
19. B. J. Pettis, "Differentiation in Banach spaces," *Duke Mathematical Journal*, vol. 5 (1939), pp. 254-269.
20. B. J. Pettis, "On integration in vector spaces," *Transactions of the American Mathematical Society*, vol. 44 (1938), pp. 277-304.
21. R. S. Phillips, "On linear transformations," *Transactions of the American Mathematical Society*, vol. 48 (1940), pp. 516-541.
22. S. Saks, *Theory of the integral*, Warsaw, 1937.
23. J. Sirvint, "Schwache Kompaktheit in den Banachschen Räumen," *Comptes Rendus U. R. S. S.*, vol. 28 (1940), pp. 199-202.
24. V. Smulian, "Über lineare topologische Räume," *Recueil Mathématique*, vol. 49 (1940), pp. 425-444.
25. V. Smulian, "Sur les ensembles faiblement compact dans les espaces linéaires normés," *Communications de la Société Mathématique de Kharkhoff*, 4, vol. 14 (1937), pp. 239-242.

CONTINUA OF FINITE LINEAR MEASURE I.*

By SAMUEL EILENBERG and O. G. HARROLD, JR.

1. The α-dimensional measures. For a metric space X and numbers $\alpha \geqq 0$, $\epsilon > 0$ define[1] $L_\epsilon^\alpha(X) = \text{g. l. b.} \sum_{i=1}^{\infty} [\delta(X_i)]^\alpha$ for all decompositions $X = X_1 + X_2 + \cdots, \delta(X_i) < \epsilon$. If no such decomposition exists set $L_\epsilon^\alpha(X) = +\infty$. Since L_ϵ^α is a non-decreasing function of ϵ we may define

$$L^\alpha(X) = \lim L_\epsilon^\alpha(X), \epsilon \to 0.$$

This number (which may be $+\infty$) is called the *α-dimensional measure* of X. The number $L^\alpha(A)$ is also defined for all subsets $A \subset X$ and is an outer Carathéodory measure [10, p. 53].[2]

It is agreed that $[\delta(A)]^0 = 0, 1$ according as A is vacuous or not, and hence $L^0(X)$ is the number of points in X if it is a finite set and $+\infty$ otherwise.

The value of $L^\alpha(X)$ depends strictly on the choice of the distance function ρ in X. For the sake of clarity we shall sometimes write $L^\alpha(X, \rho)$ instead of $L^\alpha(X)$.

2. Szpilrajn's theorem. The problem. The connection between α-dimensional measures and the Menger-Urysohn dimension of X was studied by Szpilrajn [11, 7] who, among other results, has proved that a separable metric space X has a dimension $\leqq n$ if and only if there is a metrization ρ of X such that $L^{n+1}(X, \rho) = 0$. This implies that if $\dim X = n$, $L^n(X, \rho) > 0$ for every metrization ρ of X.

If X_0 is a compact metric space consisting of a countably infinite set of points and X_n is the cartesian product $I^n \times X_0$, where I^n is the n-dimensional cube, it is easily seen that $\dim X_n = n$ but $L^n(X_n, \rho) = +\infty$ for every metrization ρ. This suggests the following

PROBLEM. *To characterize a separable metric space X of dimension n for which there is a metrization ρ such that $L^n(X, \rho) < +\infty$.*

* Received July 31, 1941.

[1] For $A \subset X$ we define $\delta(A) = \text{l. u. b. } \rho(x, y)$, $x, y \in A$.

[2] The numbers in brackets refer to the bibliography at the end of the paper.

3. The integral theorem. A necessary condition which the space must satisfy can be easily obtained from the following [3]

THEOREM 1. *Given a metric space X and a real valued function $f(x)$ on X such that $|f(x_1) - f(x_2)| \leqq \rho(x_1, x_2)$ for all $x_1, x_2 \in X$, then*

$$\int_{-\infty}^{+\infty} L^\alpha[f^{-1}(t)]dt \leqq L^{\alpha+1}(X), \qquad \alpha \geqq 0,$$

where the integral is to be understood as an upper integral in the sense of Lebesgue.[4]

Since all the quantities involved are non-negative it follows that if $L^{\alpha+1}(X) < +\infty$ (or $= 0$) then $L^\alpha[f^{-1}(t)] < +\infty$ (or $= 0$) for almost all t. In particular if $L^1(X) < +\infty$, then $f^{-1}(t)$ is a finite set for almost all t.

If x_0 is a fixed point of X and $f(x) = \rho(x, x_0)$, we get the following

COROLLARY. *If $L^1(X, \rho) < \infty$ and $x_0 \in X$ then the set*

$$S(t) = \{x \mid \rho(x_0, x) = t\}$$

is finite for almost all t.

Using the function $f(x) = \rho(x_0, x)$ we also easily deduce from Theorem 1 that if $L^{n+1}(X) = 0$ then $\dim X \leqq n$, which is one half of the theorem of Szpilrajn mentioned above.

4. Rectifiable continua. In this paper attention is restricted to the case $n = 1$ and X is a continuum (i. e. a metric, compact and connected space containing at least two points). For this case the problem will be completely solved (see Section **6**). This case, the authors believe, is of particular importance on account of its connection with the problem of rectification of continuous curves. A continuum X is called rectifiable if there is a continuous mapping, $f(I) = X$, of the closed unit interval I, $0 \leqq t \leqq 1$, such that the path-length $S(f)$ is finite.[5] Since the inequality $L^1(X) \leqq S(f)$ is readily established from the definitions, it follows that, if X is rectifiable, the linear measure of X is finite. The converse of this is a consequence of the following

---

[3] This is a generalization of a theorem of Eilenberg [2].

[4] We understand the upper Lebesgue integral of a function $f \geqq 0$ to be the g. l. b. of Lebesgue integrals of all measurable functions $g \geqq f$.

[5] For a definition of $S(f)$ and simple properties of this functional see Saks [10, p. 121].

THEOREM 2. *If $L^1(X) < \infty$ there is a continuous mapping $f(I) = X$ such that $S(f) \leqq 2 \cdot L^1(X) - \delta(X)$.*

This is a generalization of a theorem of Ważewski [12] (see also [8] and [9]).

5. Continua of finite degree. The point p of the continuum X is said to be of *finite degree* provided that to each $\epsilon > 0$ there is an uncountable family of neighborhoods $\{U_\alpha\}$ of p such that (a) $\delta(U_\alpha) < \epsilon$, (b) for arbitrary U_α, U_β in $\{U_\alpha\}$ either $\bar{U}_\alpha \subset U_\beta$ or $\bar{U}_\beta \subset U_\alpha$ and (c) each $\bar{U}_\alpha - U_\alpha = F(U_\alpha)$ is a finite set. If each point of X is of finite degree, the continuum X is of *finite degree*. A continuum of finite degree is already regular in the Menger-Urysohn sense.

The point p of the continuum X is called a *local separating point* provided that some pair of points, of the component C containing p of an arbitrary neighborhood R of p, are separated by p in R. The development of these concepts is due to G. T. Whyburn who has shown that a continuum X is of finite degree if and only if every subcontinuum of X contains uncountably many local separating points of X [13]. This theorem of Whyburn implies that if X is of finite degree at all points except perhaps those of a compact 0-dimensional set it is of finite degree.

6. The main theorem. Denote by I the unit interval $0 \leqq t \leqq 1$. If $f(X) = Y$, denote by $\Phi(f)$ the set of all points in Y such that $f^{-1}(y)$ is a finite set. We come now to the principal

THEOREM 3. *For every continuum X the following conditions are equivalent:* [6]

(A) *X can be imbedded in the Hilbert cube Q_ω so as to have a finite linear measure.*

(B) *X has a homeomorphic image of finite linear measure.*

(C) *X is of finite degree.*

(D) *Given any two disjoint, closed subsets X_0, X_1 in X there is a continuous mapping $f(X) = I$ such that*

(d_1) *$f(x) = i$, for $x \in X_i$, $i = 0, 1$;*

(d_2) *$\Phi(f)$ is uncountable.*

---

[6] For the equivalence of (A), (B) and (D) see [3]. For the equivalence of (B) and (C), originally established by other methods, see abstract 45-9-321 of the *Bulletin of the American Mathematical Society*, vol. 45 (1939), p. 668.

(E) *Given any two closed, disjoint subsets X_0, X_1 in X there is a continuous mapping $f(X) = I$ such that*

(e_1) *$f(x) = i$, for $x \in X_i$, $i = 0, 1$;*

(e_2) *for every irrational $t \in I$, $t \in \Phi(f)$.*

(F) *Every subcontinuum of X contains uncountably many local separating points of X.*

(G) *X is locally connected and for every pair of closed, disjoint subsets X_0, X_1 in X there is a finite collection of disjoint, perfect sets $N_1, N_2, \cdots N_k$ such that any continuum in X intersecting both X_0 and X_1 contains some N_i.*

(H) *Given a sequence $X_0, X_1, X_2, \cdots$ of subcontinua such that $\lim X_i = X_0$ there is an integer n for which $\prod_n^\infty X_i$ is uncountable.*

The problem, as to whether or not the 3-dimensional cube, Q^3, can be substituted for the Hilbert cube in (A) is open.

In applications the condition (F) proves most useful in deciding whether or not a given continuum has a homeomorph of finite length.

It is evident that (A) $\rightarrow$ (B). The implication (B) $\rightarrow$ (C) is a consequence of the Corollary to Theorem 1. The theorem of Whyburn mentioned in Section **5** states the equivalence of (C) and (F). The equivalence of (C), (G) and (H) has been proved elsewhere [5]. Hence only the implications

$$(C) \rightarrow (D) \rightarrow (E) \rightarrow (A)$$

need to be proved. These proofs as well as the proofs of Theorem 1 and Theorem 2 are given in the following sections.

7. Proof of Theorem 1. Given a subset A of X define

$$d(A, t) = 1 \text{ if } t \in \overline{f(A)}; \qquad d(A, t) = 0 \text{ if } t \text{ non} \in \overline{f(A)}.$$

Let now

$$X = A_1^n + A_2^n + \cdots \qquad \delta(A_i^n) < 1/n$$

be a sequence of decompositions such that

$$\lim_{n\to\infty} \sum_{i=1}^{\infty} \{\delta(A_i^n)\}^{a+1} = L^{a+1}(X). \tag{7.1}$$

Define

$$d(t) = \liminf_{n\to\infty} \sum_{i=1}^{\infty} \{\delta(A_i^n)\}^a d(A_i^n, t). \tag{7.2}$$

Since $d(A_i^n, t)$ is a measurable function of t,

(7.3) $d(t)$ is measurable.

In view of the definition of L^α we have

$$L^\alpha[f^{-1}(t)] \leqq \liminf_{n\to\infty} \sum_{i=1}^{\infty} \{\delta[A_i^n \cdot f^{-1}(t)]\}^\alpha$$

But since

$$\{\delta[A_i^n \cdot f^{-1}(t)]\}^\alpha = \{\delta[A_i^n \cdot f^{-1}(t)]\}^\alpha d(A_i^n, t) \leqq \{\delta(A_i^n)\}^\alpha d(A_i^n, t)$$

it follows that

(7.4) $$L^\alpha[f^{-1}(t)] \leqq d(t).$$

From (7.2) and the lemma of Fatou [10, p. 29] it follows that

$$\int_{-\infty}^{\infty} d(t)dt \leqq \liminf_{n\to\infty} \sum_{i=1}^{\infty} \{\delta(A_i^n)\}^\alpha \int_{-\infty}^{\infty} d(A_i^n, t)dt.$$

But since

$$\int_{-\infty}^{\infty} d(A_i^n, t)dt = \text{meas}\, \overline{f(A_i^n)} \leqq \delta[f(A_i^n)] \leqq \delta(A_i^n)$$

it follows that

$$\int_{-\infty}^{\infty} d(t)dt \leqq \liminf_{n\to\infty} \sum_{i=1}^{\infty} \{\delta(A_i^n)\}^{\alpha+1}.$$

This implies

(7.5) $$\int_{-\infty}^{\infty} d(t)dt \leqq L^{\alpha+1}(X)$$

because of (7.1). The theorem is a consequence of (7.3), (7.4) and (7.5).

Notice that the case $\alpha = 0$ needs no special proof provided we understand that $\{\delta(A)\}^0 = 0$ if $A = 0$ and that $\{\delta(A)\}^0 = 1$ if $A \neq 0$.

8. Proof of Theorem 2. We shall first establish the more general

THEOREM 2a. *Let X be a continuum such that $L^1(X) < \infty$ and let $A \subset X$ be an arc with endpoints a_0, a_1. There is a mapping $f(I) = X$ such that $f(0) = a_0$, $f(1) = a_1$ and $S(f) \leqq 2L^1(X) - L^1(A)$.*

Proof. In case X is a tree (i. e. a connected graph containing no closed curve) the theorem can be established by an easy induction with respect to the number of ramification points of X.

Since $L^1(X) < \infty$, X is a locally connected continuum and therefore there is a sequence of trees $\{T_i\}$ such that

$$A \subset T_i \subset X; \qquad \text{Lim}\, T_i = X.$$

Hence we obtain a sequence of mappings $f_i(I) = T_i$ such that $f_i(0) = a_0$, $f_i(1) = a_1$ and

$$S(f_i) \leqq 2L^1(T_i) - L^1(A) \leqq 2L^1(X) - L^1(A).$$

Since the functions $\{f_i\}$ have a uniformly bounded variation there is a subsequence $\{f_{i_k}\}$ converging uniformly to a mapping $f(I) \subset X$ which clearly will satisfy all the requirements of Theorem 2a.

To prove Theorem 2 we notice that since $L(X) < \infty$, X is arcwise connected and we can chose the arc A so that $\delta(A) = \delta(X)$. Since $\delta(A) \leqq L^1(A)$ we have $2L^1(X) - L^1(A) \leqq 2L^1(X) - \delta(X)$.

9. Proof that (C) $\rightarrow$ (D). Let X_0 and X_1 be two disjoint closed subsets of X. Identifying all the points belonging to X_0 one obtains a continuum X^* which is a continuous image of X. If X is of finite degree then X^* is of finite degree at every point except perhaps the point x_0 which is the image of X_0. As we have remarked in Section **5**, X^* must therefore be of finite degree also at x_0. This implies the existence of an uncountable family $\{U_\alpha\}$ of open subsets of X such that: $X_0 \subset U_\alpha$, $U_\alpha \cdot X_1 = 0$, $\bar{U}_\alpha - U_\alpha$ is finite, either $U_{\alpha_1} \subset U_{\alpha_2}$ or $U_{\alpha_2} \subset U_{\alpha_1}$.

By a theorem of Kuratowski [7] a continuous mapping $f(X) = I$ can be constructed such that $f(x) = 0$ for $x \in X_0$, $f(x) = 1$ for $x \in X_1$ and that for all indices α, except for a countable number of them, the equality $\bar{U}_\alpha - U_\alpha = f^{-1}(t)$ holds for some t. Consequently $\Phi(f)$ is uncountable and (D) is established.

10. Proof that (D) $\rightarrow$ (E). We shall use the following

LEMMA 1. *Given a perfect set $N \subset I$ there is a continuous mapping $g(I) = I$ such that $g(0) = 0$, $g(1) = 1$, and for every irrational $t \in I$ the set $g^{-1}(t)$ is a single point of N.*

We omit the proof.

Let now f be a mapping given by (D). Since $\Phi(f)$ is uncountable and $\Phi(f)$ is a Borel set (in fact, a $G_{\delta\sigma}$) there is a perfect set $N \subset \Phi(f)$. Let $g(I) = I$ be the mapping given by Lemma 1, and define $f_1 = gf$. For every irrational t, we have $g^{-1}(t) \in N \subset \Phi(f)$ hence $t \in \Phi(f_1)$. This proves (E).

11. Regular decompositions. A decomposition

$$X = F_1 + F_2 + \cdots + F_k$$

---

[7] C. Kuratowski, "Sur les familles monotones d'ensembles fermés et leurs applications à la théorie des espaces connexes," *Fundamenta Mathematicae*, vol. 30 (1938), pp. 17-33.

will be called regular if (a) F_i is a continuum for $i = 1, 2, \cdots, k$ and (b) $F_i \cdot F_j$ is finite for $i \neq j$.

LEMMA 2. *A continuum X is regular in the Menger-Urysohn sense if and only if for every $\epsilon > 0$ there is a regular decomposition of mesh*[8] ϵ.

Proof. It is obvious that if the condition of Lemma 2 is satisfied then X is regular. Assume now that X is regular. There is then a decomposition

$$X = U_1 + U_2 + \cdots + U_m$$

where U_i is open, $\delta(U_i) < \epsilon/2$ and $B(U_i) = \bar{U}_i - U_i$ is finite for $i = 1, 2, \cdots, m$. Consider the finite set

$$B = \sum_{i=1}^{m} B(U_i) = (x_1, x_2, \cdots, x_r)$$

It is clear that every component of $X - B$ has diameter $< \epsilon/2$ and only a finite number of components of $X - B$ have a boundary consisting of more than one point. Hence we have a decomposition

$$X - B = D_1 + D_2 + \cdots + D_r + D_{r+1} + \cdots + D_k$$

where for $i \leqq r$, D_i is the sum of all components of $X - B$ with x_i as boundary and $D_{r+1}, \cdots, D_k$ are the remaining components. Taking $F_i = \bar{D}_i$ we obtain a regular decomposition of mesh ϵ.

12. The condition (I). In terms of regular decompositions we can formulate the following condition for a continuum X.

(I) *Given any two disjoint, closed subsets X_0, X_1 in X there is a continuous mapping $g(X) \subset I$ such that*

(i_1) $g(x) = 0$ *for* $x \in X_0$, $g(x) > 0$ *for* $x \in X_1$;

(i_2) *for every regular decomposition* $\{F_i\}$ *of* X, $\sum_i \delta[g(F_i)] \leqq 1$.

LEMMA 3. (E) $\rightarrow$ (I).

Proof. Let f be the function given by (E). Let $k(t)$ denote the number of points of $f^{-1}(t)$ if $f^{-1}(t)$ is finite and let $k(t) = \infty$ otherwise. It follows from (e_2) that $k(t) < \infty$ for every irrational $t \in I$.

We define the mapping $g(X) \subset I$ as follows

$$g(x) = \int_0^{f(x)} \frac{dt}{k(t)} .$$

[8] A decomposition will be said to be of mesh ϵ provided the diameter of each element is $< \epsilon$.

If $x \in X_0$ then $f(x) = 0$ and $g(x) = 0$. If $x \in X_1$ then $f(x) = 1$ and

$$g(x) = \int_0^1 \frac{dt}{k(t)};$$

therefore $g(x) > 0$ since $k(t) < \infty$ on a set of positive measure. Condition (i_1) is thus fulfilled.

Now let a regular decomposition $\{F_i\}$ be given. Since F_i is a continuum, $f(F_i)$ is an interval or a point:

$$f(F_i) = [a_i, b_i] \text{ where } a_i \leqq b_i.$$

It follows that

$$\delta[g(F_i)] = \int_0^{b_i} \frac{dt}{k(t)} - \int_0^{a_i} \frac{dt}{k(t)} = \int_{a_i}^{b_i} \frac{dt}{k(t)}$$

and therefore

$$\sum_i \delta[g(F_i)] = \sum_i \int_{a_i}^{b_i} \frac{dt}{k(t)} \tag{12.1}$$

Let $0 = c_0 < c_1 \cdots < c_r = 1$ be the sequence obtained by arranging the points $\{a_i\}$, $\{b_i\}$ in the natural order. Each of the intervals $[a_i, b_i]$ will be the sum of a number of consecutive intervals $[c_j, c_{j+1}]$ and therefore from (12.1) we obtain

$$\sum_i \delta[g(F_i)] = \sum_{j=0}^{r-1} k_j \int_{c_j}^{c_{j+1}} \frac{dt}{k(t)} \tag{12.2}$$

where k_j is the number of indices i for which $[c_j, c_{j+1}] \subset [a_i, b_i]$. It follows that $[c_j, c_{j+1}] \subset f(F_i)$ for k_j different indices i. Since $F_i \cdot F_j$ is finite for $i \neq j$ we deduce that $f^{-1}(t)$ contains at least k_j points for every $t \in [c_j, c_{j+1}]$ except perhaps a finite set. Hence $k_j \leqq k(t)$ for almost every $t \in [c_j, c_{j+1}]$ and therefore

$$k_j \int_{c_j}^{c_{j+1}} \frac{dt}{k(t)} \leqq c_{j+1} - c_j.$$

Using (12.2) we obtain (i_2).

13. Proof that (E) $\rightarrow$ (A). Let $U_1, U_2, \cdots$ be a sequence of open sets forming a base for X. Let $\{(X_0^n, X_1^n)\}$ be the sequence of all pairs (X_0^n, X_1^n) such that $X_0^n = \bar{U}_i$, $X_1^n = X - U_j$ and $X_0^n \cdot X_1^n = 0$.

Since (E) implies (I), for every pair (X_0^n, X_1^n) we may consider a function g_n satisfying (I). Given $x \in X$ the sequence

$$h(x) = \{g_1(x), 2^{-2}g_2(x), \cdots, n^{-2}g_n(x), \cdots\}$$

is a point of the Hilbert cube Q_ω. Clearly the mapping $h(X) \subset Q_\omega$ is continuous.

If $x_0, x_1 \in X$ and $x_0 \neq x_1$ then for some index n we have $x_0 \in X_0^n$, $x_1 \in X_1^n$ and because of (i_1), $g_n(x_0) < g_n(x_1)$ and consequently $h(x_0) \neq h(x_1)$. Hence h is a homeomorphism.

Since for $x_0, x_1 \in X$

$$\rho[h(x_0), h(x_1)] = \{\sum_{n=1}^{\infty} n^{-4} \mid g_n(x_0) - g_n(x_1) \mid^2\}^{\frac{1}{2}}$$

it follows that

$$\delta[h(A)] \leqq \{\sum_{n=1}^{\infty} n^{-4}(\delta[g_n(A)])^2\}^{\frac{1}{2}} \leqq \sum_{n=1}^{\infty} n^{-2}\,\delta[g_n(A)]$$

for every subset A of X.

Given a regular decomposition $\{F_i\}$ of X we have

$$\sum_i \delta[h(F_i)] \leqq \sum_i \sum_{n=1}^{\infty} n^{-2}\delta[g_n(F_i)] \leqq \sum_{n=1}^{\infty} n^{-2} \sum_i \delta[g_n(F_i)]$$

and since $\sum_i \delta[g_n(F_i)] \leqq 1$ because of (i_2), we obtain

$$\sum_i \delta[h(F_i)] \leqq \sum_{n=1}^{\infty} n^{-2}. \tag{13.1}$$

Now, since X satisfies (E), it is clear that X is regular and therefore in view of Lemma 2 there are regular decompositions of X of arbitrarily small mesh. Hence (13.1) implies that

$$L^1[h(X)] \leqq \sum_{n=1}^{\infty} n^{-2} < \infty.$$

This concludes the proof of Theorem 3. Notice that in proving (A) we have used (I) and the fact that X is regular. Since (E) implies (I) we therefore obtain

THEOREM 3a. *For a continuum X regular in the sense of Menger-Urysohn conditions* (A) — (H) *and* (I) *are equivalent.*

UNIVERSITY OF MICHIGAN.
LOUISIANA STATE UNIVERSITY.

BIBLIOGRAPHY

1. Damköhler, W., "Struktur der geschlossenen rectifizierbaren Kurven," *Journal für die reine und angewandte Mathematik*, vol. 177 (1937), pp. 37-54.
2. Eilenberg, S., "On Φ measures," *Annales de la Société Polonaise de Mathématique*, vol. 17 (1938), pp. 252-253.

3. ———, "On continua of finite length," *Annales de la Société Polonaise de Mathématique*, vol. 17 (1938), pp. 253-254.
4. Favard, J., "Sur la structure des continus rectifiables," *Journal de Mathématiques pures et appliquées*, vol. 15 (1936), pp. 321-332.
5. Harrold, O. G., "Continua of finite degree and certain product sets," *Bulletin of the American Mathematical Society*, vol. 46 (1940), pp. 951-953.
6. Hausdorff, F., "Dimension und äusseres Mass," *Mathematische Annalen*, vol. 79 (1919), pp. 157-179.
7. Hurewicz, W. and Wallman, H., *Dimension theory, Princeton Mathematical Series*, vol. 4 (Princeton), 1941.
8. Kolmogoroff, A., "Beiträge zur Masstheorie," *Mathematische Annalen*, vol. 107 (1933), pp. 351-366.
9. Milgram, A. N., "On shortest paths through a set," *Reports of a mathematical Colloquium*, second series 2 (1940), pp. 39-44.
10. Saks, S., *Theory of the integral, Monografje Matematyczne*, vol. 7, Warszawa (1937).
11. Szpilrajn, E., "La dimension et la mesure," *Fundamenta Mathematicae*, vol. 28 (1937), pp. 81-89.
12. Ważewski, T., "Kontinua prostowalne w związku z funkcjami i odwzorowaniami absolutnie ciągłemi," *Annales de la Société Polonaise de Mathématique*, supplement, vol. 3 (1937), pp. 9-49.
13. Whyburn, G. T., "Concerning continua of finite degree and local separating points," *American Journal of Mathematics*, vol. 57 (1935), pp. 11-18.

A FUNCTION-THEORETIC IDENTITY.*

By D. C. SPENCER.

1. Let $n(r, w)$ denote the number of roots of $f = w$ in $|z| < r$. We have then the following rather general result:

THEOREM 1. *Suppose that $g(R)$, defined for $R \geq 0$, is absolutely continuous*[1] *and that*

$$G(R) = \int g(R)\, d(\ln R). \tag{1.1}$$

Then for any f regular in $|z| \leq r$ we have

$$r\frac{d}{dr}\int_{-\pi}^{\pi} G(|f(re^{i\phi})|)\, d\phi = \int_0^r \int_{-\pi}^{\pi} g'(|f|)\cdot|f|^{-1}\cdot|f'|^2 \rho\, d\rho\, d\phi + 2\pi g(0)\cdot n(r,0) \tag{1.2}$$

(except in cases when $g(0) \neq 0$ and $f = 0$ on $|z| = r$).

Since the derivative appears on the left-hand side of (1.2), the additive constant implicit in the definition of $G(R)$ does not matter.

For functions meromorphic in $|z| \leq r$, the identity becomes

$$r\frac{d}{dr}\int_{-\pi}^{\pi} G(|f(re^{i\phi})|)\, d\phi = \int_0^r \int_{-\pi}^{\pi} g'(|f|)\cdot|f|^{-1}\cdot|f'|^2 \rho\, d\rho\, d\phi + 2\pi g(0)\, n(r,0) - 2\pi g(\infty)\, n(r,\infty).$$

This can be seen almost at once from the argument given in Section **5**, the only modification (which is rather trivial) being in (5.3) of Theorem 6. For convenience we suppose throughout that f is regular[2] in $|z| < 1$, and that $0 < |z| = r < 1$.

Taking $g(R) = 1$, so that $g'(R) = 0$, $G(R) = \ln R$, we obtain

$$r\frac{d}{dr}\int_{-\pi}^{\pi} \ln|f|\, d\phi = 2\pi n(r,0), \tag{1.3}$$

* Received August 24, 1941.

[1] Absolute continuity on the set of values of $|f|$ is of course sufficient.

[2] There is no loss of generality whatever in considering regular functions; the method given here is applicable with trivial modification to meromorphic functions.

which is Jensen's formula. Next if $\lambda > 0$ [3], the choice $g(R) = \lambda R^\lambda$ gives

$$r\frac{d}{dr}\int_{-\pi}^{\pi} |f|^\lambda d\phi = \lambda^2 \int_0^r \int_{-\pi}^{\pi} |f|^{\lambda-2} \cdot |f'|^2 \rho d\rho d\phi, \tag{1.4}$$

the identity of Hardy-Stein. Thirdly, if $f \not\equiv 0$, we may take $g(R) = \ln R$, obtaining

$$r\frac{d}{dr}\int_{-\pi}^{\pi} \ln^2 |f| \, d\phi = 2\int_0^r \int_{-\pi}^{\pi} \left|\frac{d}{dz} \ln f\right|^2 \rho d\rho d\phi. \tag{1.5}$$

This formula may also be found by dividing both sides of (1.4) by λ^2 and letting $\lambda \to 0$.

Taking $g(R) = R^2/(1+R^2)$, $G(R) = \ln[(1+|a|^2)(1+R^2)]^{\frac{1}{2}}$, we have (since $g(0) = 0$)

$$r\frac{d}{dr}\int_{-\pi}^{\pi} \ln[(1+|a|^2)(1+|f|^2)]^{\frac{1}{2}} d\phi = 2\int_0^r \int_{-\pi}^{\pi} \frac{|f'|^2}{(1+|f|^2)^2} \rho d\rho d\phi. \tag{1.6}$$

Replacing f by $f - a$ in (1.3), subtracting the resulting identity from (1.6) and then integrating, we obtain, in the notation of Ahlfors,[4]

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \ln \frac{1}{[f,a]} \, d\phi + N(r,a) = \int_0^r \frac{S(\rho)}{\rho} \, d\rho. \tag{1.7}$$

This is Ahlfors' formulation of Nevanlinna's first main theorem; similar identities were found earlier by Shimizu (3) and Ahlfors (1). Both authors use methods quite different from that given here.

Finally suppose that $\lambda > 0$, and take $g(R) = \text{Min}(\lambda R^\lambda, \lambda)$. Then if $f(0) = 0$ we obtain, after integrating, the following curious identity,

$$\lambda\int_{-\pi}^{\pi} \ln^+ |f| \, d\phi = \lambda^2 \int_0^r \left\{ \int_0^\rho \int_{\substack{-\pi \\ |f|\leq 1}}^{\pi} |f|^{\lambda-2} |f'|^2 \rho_1 d\rho_1 d\phi \right\} \frac{d\rho}{\rho} - \int_{\substack{-\pi \\ |f|\leq 1}}^{\pi} |f|^\lambda d\phi, \tag{1.8}$$

where $\ln^+ |f| = \text{Max}(\ln |f|, 0)$, and where the integration on the right-hand side is extended only over values z for which $|f| \leq 1$. Integration by ϕ alone refers (as above) to the circumference $|z| = r$.

2. We write

$$A(r) = A_g(r) = \int_0^r \int_{-\pi}^{\pi} g'(|f|) \cdot |f|^{-1} \cdot |f'|^2 \rho d\rho d\phi. \tag{2.1}$$

$A(r)$, an area-functional, may be given the following geometrical interpre-

---

[3] If $f \not\equiv 0$ in $|z| \leqq r$, (1.4) is true for all real λ.

[4] Ahlfors (2). $S(r)$ is the area, divided by π, of the stereographic projection on the Riemann sphere (of diameter unity) of the transform of $|z| < r$ by f; and $[a, b]$ is the straight-line distance between the projections of a and b on this sphere.

tation. We suppose first that $g(R)$ is non-negative and non-decreasing. Let $W(r)$ denote the map of $|z|<r$ by f, $w=Re^{i\Phi}$ a point of $W(r)$, and let $V(r)=V_g(r)$ be the aggregate of points $v=[2g(R)]^{\frac{1}{2}}\cdot e^{i\Phi}$ for which $Re^{i\Phi}\subset W(r)$. Then $V(r)$ is a uniform distortion of $W(r)$ in the radial direction, and $A(r)$ is the area of $V(r)$, multiply covered points being counted multiply. In general, since $g(R)$ is absolutely continuous, we may write

$$g(R)=g_1(R)-g_2(R), \tag{2.2}$$

where g_1 and g_2 are each non-negative, non-decreasing, and absolutely continuous; and

$$A(r)=A_1(r)-A_2(r), \tag{2.3}$$

where A_1 and A_2 are the areas of V_{g_1} and V_{g_2} respectively.

3. We define

$$p(r,R)=\frac{1}{2\pi}\int_{-\pi}^{\pi}n(r,Re^{i\Phi})d\Phi, \tag{3.1}$$

$$n(w)=n(1,w)=\lim_{r\to 1}n(r,w), \tag{3.2}$$

$$p(R)=p(1,R)=\lim_{r\to 1}p(r,R)=\frac{1}{2\pi}\int_{-\pi}^{\pi}n(Re^{i\Phi})d\Phi, \tag{3.3}$$

the limits (though perhaps infinite) existing in virtue of the fact that $n(r,Re^{i\Phi})$, and so $p(r,R)$, are non-decreasing functions of r, $0<r<1$.

If

$$\int_0^{R_1}p(R)d[\psi(R)]\leq \boldsymbol{p}\int_0^{R_1}d[\psi(R)] \tag{3.4}$$

for all $R_1>0$, where $\boldsymbol{p}$ is a positive (though not necessarily integral) number, we say that f is mean $\boldsymbol{p}$-valent (ψ). Here we suppose throughout that ψ is an absolutely continuous, non-decreasing function, and that $\psi(0)=0$. Then (by absolute continuity) (3.4) becomes

$$\int_0^{R_1}p(R)d[\psi(R)]=\int_0^{R_1}p(R)\cdot\psi'(R)dR\leq \boldsymbol{p}\psi(R_1)\qquad(R_1>0). \tag{3.5}$$

In particular, $\boldsymbol{p}$-valent functions are mean $\boldsymbol{p}$-valent (ψ). The special case of (3.4) when $\psi=R^2$ is ordinary mean $\boldsymbol{p}$-valency, and this has been discussed elsewhere.[5] Roughly, (3.4) is the stronger the more rapidly ψ increases with R.

We note first that mean $\boldsymbol{p}$-valency (ψ) implies that $n(0)\leq \boldsymbol{p}$. For

[5] See Spencer (4), (5), and (7).

there is a circle of radius $\delta = \delta(f) > 0$ every point of which is covered exactly $n(0)$-times, and if $n(0) > \boldsymbol{p}$ we should have

$$(3.6) \quad \int_0^{\delta} p(R)d[\psi(R)] \geq n(0) \cdot \int_0^{\delta} d[\psi(R)] > \boldsymbol{p} \int_0^{\delta} d[\psi(R)],$$

in contradiction with (3.4). If $\boldsymbol{p} = n(0)$, we say that f is mean $n(0)$-valent (ψ). Since f is regular in $|z| < 1$, $p(\infty) = 0$, and

$$(3.7) \quad \int_0^{\infty} d[-p(R)] = p(0) - p(\infty) = p(0) = n(0).$$

Mean $n(0)$-valency may therefore be written in the form

$$(3.8) \quad \int_0^{R_1} p(R)d[\psi(R)] \leq \psi(R_1) \cdot \int_0^{\infty} d[-p(R)] \qquad (R_1 > 0).$$

Secondly, since $p(r, R) \leq p(R)$, we see that

$$(3.9) \quad \int_0^{R_1} p(r, R)d[\psi(R)] \leq \int_0^{R_1} p(R)d[\psi(R)] \qquad (0 < r < 1, R_1 > 0).$$

Therefore a mean $\boldsymbol{p}$-valent f is *a fortiori* mean $\boldsymbol{p}$-valent inside $|z| < r < 1$.

The following theorem is useful when combined with Theorem 1.

THEOREM 2. *Suppose that f is mean $\boldsymbol{p}$-valent (ψ). Then*

$$(3.10) \quad \int_0^{\infty} \phi\{\psi(R)\}d[-p(R)] \geq \boldsymbol{p}\phi\left\{\frac{1}{\boldsymbol{p}} \int_0^{\infty} \psi(R)d[-p(R)]\right\}$$

if ϕ is convex (that is, if $\phi'' \geq 0$) and if $\phi(0) \leq 0$; and

$$(3.11) \quad \int_0^{\infty} \phi\{\psi(R)\}d[-p(R)] \leq \boldsymbol{p}\phi\left\{\frac{1}{\boldsymbol{p}} \int_0^{\infty} \psi(R)d[-p(R)]\right\}$$

if ϕ is concave (that is, if $\phi'' \leq 0$) and if $\phi(0) \geq 0$.

When f is mean $n(0)$-valent (ψ), we may replace $\boldsymbol{p}$ by $n(0) = \int_0^{\infty} d[-p(R)]$ in (3.10) and (3.11), and these inequalities assume a more elegant form.

The restrictions that $\phi(0)$ be non-positive and non-negative respectively in (3.10) and (3.11) arise from the fact that a mean $\boldsymbol{p}$-valent function f is *a fortiori* mean $\boldsymbol{p}_1$-valent if $\boldsymbol{p}_1 > \boldsymbol{p}$, and may therefore be dropped when f is mean $n(0)$-valent and $\boldsymbol{p}$ is replaced by $n(0)$. To see the necessity of these conditions we have only to take $\phi(R) = R + \delta$, δ positive in (3.10), negative in (3.11), to obtain contradictions of both if $\boldsymbol{p}$ is large enough.

Choosing now $\psi(R) = R^2$, $\phi(R) = g([R]^{\frac{1}{2}})$, so that mean $\boldsymbol{p}$-valency (ψ)

becomes ordinary mean $\boldsymbol{p}$-valency, and $g(R) = \phi\{\psi(R)\}$, we have, on integrating by parts,

$$(3.12)\qquad \int_0^\infty \phi\{\psi(R)\}d[-p(r,R)] = \int_0^\infty p(r,R)d[\phi\{\psi(R)\}] + \phi\{\psi(0)\}\cdot n(r,0)$$

since $p(\infty) = 0$;

$$= \int_0^\infty p(r,R)d[g(R)] + g(0)\cdot n(r,0)$$

$$= \frac{1}{2\pi}\int_0^\infty \int_{-\pi}^{\pi} n(r, Re^{i\Phi})\cdot \frac{g'(R)}{R}\, RdRd\Phi + g(0)\cdot n(r,0)$$

$$= \frac{1}{2\pi}\int_0^r \int_{-\pi}^{\pi} g'(|f|)\cdot |f|^{-1}\cdot |f'|^2 \rho d\rho d\phi + g(0)\cdot n(r,0)$$

$$= r\frac{d}{dr}\cdot\frac{1}{2\pi}\int_{-\pi}^{\pi} G(|f|)d\phi$$

by Theorem 1. Finally

$$(3.13)\qquad \int_0^\infty \psi(R)d[-p(r,R)] = \int_0^\infty R^2 d[-p(r,R)] = \int_0^\infty p(r,R)d(R^2),$$

and so, by Theorem 2, we obtain

THEOREM 3. *Suppose that f is mean $\boldsymbol{p}$-valent. Then*

$$(3.14)\qquad r\frac{d}{dr}\cdot\frac{1}{2\pi}\int_{-\pi}^{\pi} G(|f|)d\phi \geq \boldsymbol{p}g\left\{\left[\frac{1}{\boldsymbol{p}}\int_0^\infty p(r,R)d(R^2)\right]^{\frac{1}{\boldsymbol{p}}}\right\}$$

if $g([R]^{\frac{1}{\boldsymbol{p}}})$ is convex and $g(0) \leq 0$;

$$(3.15)\qquad r\frac{d}{dr}\cdot\frac{1}{2\pi}\int_{-\pi}^{\pi} G(|f|)d\phi \leq \boldsymbol{p}g\left\{\left[\frac{1}{\boldsymbol{p}}\int_0^\infty p(r,R)d(R^2)\right]^{\frac{1}{\boldsymbol{p}}}\right\}$$

if $g([R]^{\frac{1}{\boldsymbol{p}}})$ is concave and $g(0) \geq 0$.

If in addition $f(0) = 0$, $g(0) = 0$, and $G(0) = 0$, we obtain by integrating (3.14) and (3.15) from 0 to r, $0 < r < 1$, the following result:

THEOREM 4. *Suppose that f is mean $\boldsymbol{p}$-valent, and that $f(0) = 0$, $g(0) = 0$, and $G(0) = 0$. Then*

$$(3.16)\qquad \frac{1}{2\pi}\int_{-\pi}^{\pi} G(|f|)d\phi \geq \boldsymbol{p}\int_0^r g\left\{\left[\frac{1}{\boldsymbol{p}}\int_0^\infty p(\rho,R)d(R^2)\right]^{\frac{1}{\boldsymbol{p}}}\right\}\frac{d\rho}{\rho}$$

if $g([R]^{\frac{1}{\boldsymbol{p}}})$ is convex;

$$(3.17)\qquad \frac{1}{2\pi}\int_{-\pi}^{\pi} G(|f|)d\phi \leq \boldsymbol{p}\int_0^r g\left\{\left[\frac{1}{\boldsymbol{p}}\int_0^\infty p(\rho,R)d(R^2)\right]^{\frac{1}{\boldsymbol{p}}}\right\}\frac{d\rho}{\rho}$$

if $g([R]^{\frac{1}{\boldsymbol{p}}})$ is concave.

Taking in particular $g = \lambda R^\lambda$, $G(R) = R^\lambda$, we have

$$(3.18)\quad \frac{1}{2\pi}\int_{-\pi}^{\pi} |f|^\lambda d\phi \leq \boldsymbol{p}^{1-(\lambda/2)}\cdot\lambda\cdot\int_0^r \left\{\int_0^\infty p(\rho,R)\,d(R^2)\right\}^{\lambda/2}\frac{d\rho}{\rho}$$
$$= \boldsymbol{p}^{1-(\lambda/2)}\cdot\lambda\cdot\int_0^r \left\{\frac{1}{\pi}\int_0^\rho\int_{-\pi}^{\pi}|f'|^2\rho_1 d\rho_1 d\phi\right\}^{\lambda/2}\frac{d\rho}{\rho}$$
$$= \boldsymbol{p}^{1-(\lambda/2)}\cdot\lambda\cdot\int_0^r \left\{\sum_{\nu=1}^{\infty}\nu\,|a_\nu|^2\rho^{2\nu}\right\}^{\lambda/2}\frac{d\rho}{\rho}$$

if $0 < \lambda \leq 2$, where

$$(3.19)\qquad f = \sum_{\nu=1}^{\infty} a_\nu z^\nu;$$

and the reverse inequality is true if $\lambda > 2$. Writing

$$(3.20)\qquad M(\rho, f) = \operatorname*{Max}_{|z|=\rho}|f|,$$

mean $\boldsymbol{p}$-valency implies that

$$(3.21)\quad \int_0^\infty p(\rho,R)\,d(R^2) = \int_0^{M(\rho,f)} p(\rho,R)\,d(R^2) \leq \int_0^{M(\rho,f)} p(R)\,d(R^2) \leq \boldsymbol{p}M^2(\rho,f),$$

and substitution of (3.21) into (3.18) yields the weaker result (proved in my paper (4))

$$(3.22)\qquad \frac{1}{2\pi}\int_{-\pi}^{\pi}|f|^\lambda d\phi \leq \boldsymbol{p}\lambda\int_0^r M^\lambda(\rho,f)\frac{d\rho}{\rho} \qquad (0<\lambda\leq 2).$$

The significance of these inequalities becomes clearer when they are expressed in terms of integrals taken over the boundary of $W(r)$, the transform of $|z| < r$ by f. For simplicity we consider (3.18). Taking $g = \lambda R^\lambda$, $G = R^\lambda$, in formula (5.1) below, we have

$$(3.23)\qquad r\frac{d}{dr}\int_{-\pi}^{\pi}|f|^\lambda d\phi = \lambda\int_{|z|=r} R^\lambda d\Phi,$$

where $f(re^{i\phi}) = Re^{i\Phi}$. If now

$$(3.24)\qquad d\Phi \geq 0$$

for $0 < r < 1$, we have for $0 < \lambda \leq 2$, by Hölder's inequality applied to the right-hand side of (3.23),

$$(3.25)\quad r\frac{d}{dr}\int_{-\pi}^{\pi}|f|^\lambda d\phi \leqq \lambda\Big\{\int_{|z|=r} d\Phi\Big\}^{1-(\lambda/2)}\cdot\Big\{\int_{|z|=r} R^2 d\Phi\Big\}^{\lambda/2}$$
$$= \lambda\{2\pi n(r,0)\}^{1-(\lambda/2)}\cdot\Big\{\int_{|z|=r} R^2 d\Phi\Big\}^{\lambda/2}$$

by the argument principle, if $f \neq 0$ on $|z| = r$. But, by formula (5.3) below with $g = R^2$,

$$\int_{|z|=r} R^2 d\Phi = 2\pi \int_0^\infty p(r,R)\, d(R^2). \tag{3.26}$$

Substituting from (3.26) into (3.25), and integrating from 0 to r, we obtain (3.18) (with p replaced by $n(r,0)$). Similarly if $\lambda > 2$ we have the reverse inequality.

But (3.24) is much stronger than mean p-valency, and is not satisfied even by *schlicht* functions. We see, however, that integrals taken over the boundaries of mean p-valent domains behave like integrals in which $d\Phi \geq 0$.

4. The final result of the paper lies in a somewhat different direction, and is an application of the Ahlfors-Shimizu identity (1.7) in which we obtain a localization of the following well-known theorem of Nevanlinna:

(N) *If*

$$T(r) = O(1) \qquad (0 < r < 1), \tag{4.1}$$

where $T(r)$ is Nevanlinna's characteristic function, then for almost all θ, f tends to a limit as z tends non-tangentially to $e^{i\theta}$.

We suppose that $0 < \alpha < \pi/2$, denote by $\Gamma_\alpha(\theta)$ the domain defined by

$$|z| < 1, \qquad |e^{i\theta} - z| < 5/4, \qquad |\arg(e^{i\theta} - z) - \theta| < \alpha,$$

and write

$$S_\alpha(\theta) = \frac{1}{\pi} \iint_{\Gamma_\alpha(\theta)} \frac{|f'|^2}{(1+|f|^2)^2}\, r dr d\phi. \tag{4.2}$$

THEOREM 5. *Suppose that $f(z)$ is regular for $|z| < 1$, and that to each θ of a set $\mathcal{E}$, $0 < m\mathcal{E} \leq 2\pi$,[6] there is a positive number $\alpha = \alpha(\theta)$ for which $S_\alpha(\theta)$ is finite. Then for almost all θ of $\mathcal{E}$ and for every β, $0 < \beta < \pi/2$, the function $f(z)$ tends to a finite limit as $z \to e^{i\theta}$ in $\Gamma_\beta(\theta)$.*

We note that

$$S_\alpha(\theta) < \iint_{\Gamma_\alpha(\theta)} |f'|^2 r dr d\theta = A_\alpha(\theta), \tag{4.3}$$

say; hence if $A_\alpha(\theta)$ is finite in $\mathcal{E}$ the result of Theorem 4 is true *a fortiori*. Theorem 5 with the hypothesis replaced by the stronger condition that $A_\alpha(\theta)$ is finite in $\mathcal{E}$ was suggested to the author by Professor A. Zygmund.

---

[6] $m\mathcal{E}$ denotes the linear measure of $\mathcal{E}$.

Let $\bar{\Gamma}_a(\theta)$ be the subset of $\Gamma_a(\theta)$ in which $|f| \leq 1$, and write

$$\bar{A}_a(\theta) = \iint_{\bar{\Gamma}_a(\theta)} |f'|^2 r dr d\theta. \tag{4.4}$$

Using the identity (1.8) in place of (1.7) and noting that

$$\int_{-\pi}^{\pi} ln^+ |f| \, d\phi = O(1) \tag{4.5}$$

is equivalent to (4.1), a trivial modification of the proof given in Section 7 shows that the statement of Theorem 5 is true with $S_a(\theta)$ replaced by $\bar{A}_a(\theta)$.

5. We turn now to the proofs of Theorems 1, 2, and 5, and begin with Theorem 1, the proof of which is striking in its simplicity.

Theorem 1 is equivalent to the following theorem, in which we write $f = Re^{i\Phi}$:

THEOREM 6.

$$\int_{|z|=r} g(R) d\Phi = r \frac{d}{dr} \int_{-\pi}^{\pi} G(|f(re^{i\phi})|) d\phi; \tag{5.1}$$

$$\int_0^r \int_{-\pi}^{\pi} g'(|f|) \cdot |f|^{-1} \cdot |f'|^2 \rho d\rho d\phi = 2\pi \int_0^{\infty} p(r, R) d\{g(R)\}; \tag{5.2}$$

$$\int_{|z|=r} g(R) d\Phi = 2\pi \int_0^{\infty} p(r, R) d\{g(R)\} + 2\pi g(0) \cdot n(r, 0). \tag{5.3}$$

We have given elsewhere [7] a proof of the special case of Theorem 6 when $g = \lambda R^{\lambda}$, and the idea here is similar. The main difference is that we must now consider the possibility that $g(0) \neq 0$.

We suppose that $f \neq 0$ on $|z| = r$, and we recall again that, by the absolute continuity of g,

$$\int_0^{\infty} p(r, R) d\{g(R)\} = \int_0^{\infty} p(r, R) \cdot g'(R) \cdot dR. \tag{5.4}$$

We begin with (5.1). Since $f \neq 0$ on $|z| = r$, $\ln f = \ln R + i\Phi$ is regular on this circumference, and we have

$$\frac{\partial(\ln R)}{\partial r} = \frac{\partial \Phi}{r \partial \phi}; \qquad \frac{\partial \Phi}{\partial \phi} = \frac{r}{R} \frac{\partial R}{\partial r}. \tag{5.5}$$

Since, differentiating both sides of (1.1), $g(R) = RG'(R)$, we have by (5.5)

[7] Spencer (6).

$$(5.6)\qquad \int g(R)\,d\Phi = \int g(R)\,\frac{\partial\Phi}{\partial\phi}\,d\phi$$
$$= \int g(R)\,\frac{r}{R}\,\frac{\partial R}{\partial r}\,d\phi = r\,\frac{d}{dr}\int_{-\pi}^{\pi} G(R)\,d\phi,$$

and this is (5.1).

Next,

$$(5.7)\qquad \int_0^r \int_{-\pi}^{\pi} g'(|f|)\cdot|f|^{-1}\cdot|f'|^2\rho\,d\rho\,d\phi$$
$$= \int_0^{\infty}\int_{-\pi}^{\pi} n(r, Re^{i\Phi})\cdot\frac{g'(R)}{R}\cdot R\,dR\,d\Phi$$
$$= 2\pi\int_0^{\infty} p(r,R)\cdot g'(R)\cdot dR = 2\pi\int_0^{\infty} p(r,R)\,d\{g(R)\},$$

which is (5.2).

Finally, since $g(R)$ is absolutely continuous,

$$(5.8)\qquad g(R) = g_1(R) - g_2(R),$$

where g_1 and g_2 are non-negative, non-decreasing, absolutely continuous functions, and

$$(5.9)\qquad \int g(R)\,d\Phi = \int g_1(R)\,d\Phi - \int g_2(R)\,d\Phi.$$

We write

$$(5.10)\qquad \tfrac{1}{2}\underline{R}^2 = g_1(R),$$

so that

$$(5.11)\qquad \int g_1(R)\,d\Phi = \tfrac{1}{2}\int \underline{R}^2\,d\Phi.$$

As above, let $W(r)$ denote the transform of $|z|<r$ by f, and let $B(r)$ be its boundary ($B(r)$ is then the transform of $|z|=r$ by f). Also let $V(r)$ be the set of points $v=[2g_1(R)]^{\frac{1}{2}}\cdot e^{i\Phi}$ for which $Re^{i\Phi}\subset W(r)$, $\mathfrak{B}(r)$ its *external* boundary—that is to say, $\mathfrak{B}(r)$ is the curve composed of points $v=[2g_1(R)]^{\frac{1}{2}}\cdot e^{i\Phi}$, $Re^{i\Phi}\subset B(r)$. Then (See Fig. 1) $V(r)$ and $\mathfrak{B}(r)$ both lie in the annulus $[2g_1(0)]^{\frac{1}{2}}\le R\le[2g_1(\infty)]^{\frac{1}{2}}$, and $\mathfrak{B}(r)$ bounds a domain D which, if $g_1(0)\neq 0$, is composed of two parts: (i) $V(r)$; (ii) the circle C, $0\le\underline{R}<[2g_1(0)]^{\frac{1}{2}}$, every point of which is covered exactly $n(r,0)$-times since

$$(5.12)\qquad \frac{1}{2\pi}\int_{\mathfrak{B}(r)} d\Phi = \frac{1}{2\pi}\int_{|z|=r} d\Phi = n(r,0)$$

by the argument principle.

We write $A(G)$ for the area, multiply-covered points counted multiply, of a domain G. Since D is closed by the curve $w=\underline{R}e^{i\Phi}$, the right-hand side of (5.11) is the integral over this domain of a sectorial element, and therefore

$$(5.13)\qquad \int g(R)\,d\Phi = \tfrac{1}{2}\int \underline{R}^2 d\Phi = A(D) = A(V(r)) + A(C).$$

Next, $n\{r, g_1^{-1}(\frac{1}{2}\underline{R}^2)\cdot e^{i\Phi}\}$ sheets of $V(r)$ lie over the point $\underline{R}e^{i\Phi}$, and hence

$$(5.14)\qquad A(V(r)) = \int_{\substack{-\pi \\ [2g_1(0)]^{\frac{1}{2}}}}^{\substack{\pi \\ [2g_1(\infty)]^{\frac{1}{2}}}} n\{r, g_1^{-1}(\tfrac{1}{2}\underline{R}^2)\cdot e^{i\Phi}\}\,\underline{R}\,d\underline{R}\,d\Phi = 2\pi\int_0^\infty p(r,R)\,d[g_1(R)]$$

Finally

$$(5.15)\qquad A(C) = n(r,0)\cdot\pi\{[2g_1(0)]^{\frac{1}{2}}\}^2 = 2\pi g_1(0)\cdot n(r,0).$$

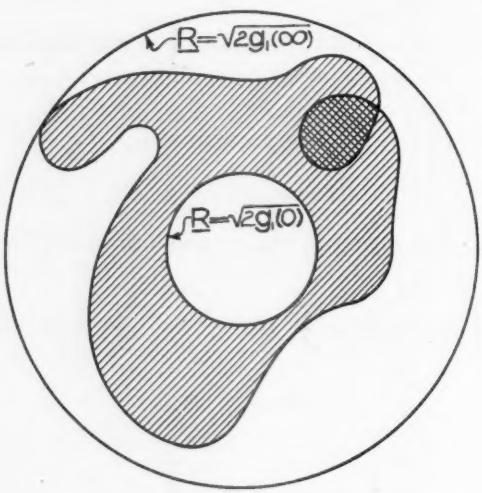


Fig. 1.

Substituting (5.14) and (5.15) into (5.13), we obtain (5.3) with g replaced by g_1. But the same formula is true for g_2, and by subtraction (using (5.8)) we obtain (5.3).

This completes the proof of Theorem 6, and so of Theorem 1.

6. We take next Theorem 2, and we may suppose that

$$(6.1)\qquad \int_0^\infty \psi(R)\,d[-p(R)] = \boldsymbol{p}\cdot\psi(R_0) < \infty.$$

For, having proved the theorem in this case, the general result follows by a trivial passage to the limit. We define

$$(6.2)\qquad p_0(R) = \begin{cases} \boldsymbol{p}, & \text{if } 0 \le R < R_0, \\ 0, & \text{otherwise,} \end{cases}$$

and then

$$(6.3)\quad \int_0^{R_1} p(R)\,d[\psi(R)] \leq \text{Min}\,\{\boldsymbol{p}\psi(R_1), \boldsymbol{p}\psi(R_0)\} = \int_0^{R_1} p_0(R)\,d[\psi(R)]$$

by the hypothesis of mean $\boldsymbol{p}$-valency and (6.1). Next, integrating twice by parts,[8]

$$(6.4)\quad \int_0^{\infty} \phi\{\psi(R)\}\,d[-p(R)] = [-p(R)\cdot\phi\{\psi(R)\}]_0^{\infty} + \int_0^{\infty} p(R)\,d[\phi\{\psi(R)\}]$$

$$= n(0)\cdot\phi(0) + \int_0^{\infty} p(R)\phi'\{\psi(R)\}\cdot\psi'(R)\,dR$$

$$= n(0)\phi(0) + [\phi'\{\psi(R)\}\cdot\int_0^{R} p(R)\,d[\psi(R)]]_0^{\infty}$$

$$-\int_0^{\infty}\{\int_0^{R_1} p(R)\,d[\psi(R)]\}\cdot\phi''\{\psi(R_1)\}\cdot\psi'(R_1)\,dR_1$$

since $\psi(0) = 0$, $p(0) = n(0)$, and $p(\infty) = 0$;

$$= n(0)\phi(0) + \phi'\{\psi(\infty)\}\cdot\boldsymbol{p}\psi(R_0)$$

$$-\int_0^{\infty}\{\int_0^{R_1} p(R)\,d[\psi(R)]\}\cdot\phi'\{\psi(R_1)\}\cdot\psi'(R_1)\,dR_1$$

by (6.1);

$$(6.5)\quad \geq \boldsymbol{p}\phi(0) + \phi'\{\psi(\infty)\}\cdot\boldsymbol{p}\psi(R_0)$$

$$-\int_0^{\infty}\{\int_0^{R_1} p_0(R)\,d[\psi(R)]\}\cdot\phi''\{\psi(R_1)\}\cdot\psi'(R_1)\,dR_1$$

if ϕ is convex and $\phi(0) \leq 0$, since $\psi' \geq 0$ (by hypothesis) and $n(0) \leq \boldsymbol{p}$;

$$= \int_0^{\infty} \phi\{\psi(R)\}\,d[-p_0(R)] = \boldsymbol{p}\phi\{\psi(R_0)\} = \boldsymbol{p}\phi\{\frac{1}{\boldsymbol{p}}\int_0^{\infty}\psi(R)\,d[-p(R)]\}$$

by (6.1) and (6.2). If ϕ is concave and $\phi(0) \geq 0$, it is plain that we have the opposite inequality.

7. It remains to prove Theorem 5. Let $\mathcal{E}_n$ be the subset of $\mathcal{E}$ for which $\alpha > 1/n$. We may suppose that $\mathcal{E}_n$ is measurable; then since $m\mathcal{E}_n \to m\mathcal{E}$ as $n \to \infty$, it is enough to prove the theorem for $\mathcal{E}_n$. In other words we may suppose that, for some fixed $\alpha > 0$, $S_\alpha(\theta)$ is finite in $\mathcal{E}$.

Next, given $\epsilon > 0$, we can find a perfect subset $\mathcal{E}_1$ of $\mathcal{E}$ such that $m(\mathcal{E} - \mathcal{E}_1) < \epsilon$, and

$$(7.1)\qquad S_\alpha(\theta) < K_1(\epsilon)$$

for θ of $\mathcal{E}$. That is to say, we may suppose that (7.1) is satisfied in $\mathcal{E}$, and that $\mathcal{E}$ is perfect.

---

[8] If $\phi'\{\psi(\infty)\} = \infty$, we suppose that $p(R) = 0$ for $M \leq R < \infty$, repeat the same argument with $\int_0^{\infty}$ replaced by $\int_0^{M}$ throughout, and then let $M \to \infty$.

We write

$$D^* = \sum_{\theta \subset \mathcal{E}} \Gamma_\alpha(\theta), \tag{7.2}$$

and Δ for the set complementary to D^* with respect to $|z| < 1$. Then Δ consists of an enumerable set of "triangular" domains Δ_ν. Let $P_1(\nu)$, $P_2(\nu)$ be the two boundary points $e^{i\theta_1}$, $e^{i\theta_2}$ of Δ_ν for which $\theta_1 \subset \mathcal{E}$, $\theta_2 \subset \mathcal{E}$, and let C_{r_0} be the circle $|z| < r_0 < 1$. We shall require the few simple facts about D^* which are summed up in the following lemma:

LEMMA 1. (i) *The distance d of the boundary of D^* from $z = 0$ satisfies*

$$d \geq \sin \alpha > 0. \tag{7.3}$$

(ii) *To any real number θ which does not belong to $\mathcal{E}$, there corresponds a $\nu = \nu(\theta)$ such that*

$$\Gamma_\alpha(\theta) = \sum_{k=1}^{4} \Gamma_k, \tag{7.4}$$

where $\Gamma_1 \subset \Delta_\nu$, $\Gamma_2 \subset C_{r_0}$, $\Gamma_3 \subset \Gamma_\alpha\{P_1(\nu)\}$ and $\Gamma_4 \subset \Gamma_\alpha\{P_2(\nu)\}$ (r_0 depending only on α).

In fact, let $P_0(\nu)$ be the point on the boundary of Δ_ν which is nearest the origin (P_0 is the "vertex" of Δ_ν), and (See Fig. 2) let ϕ be the angle P_0OP_1.

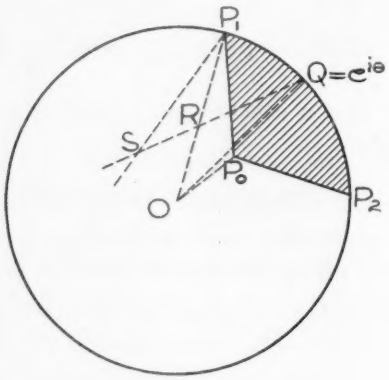


Fig. 2.

Then we have

$$OP_0 = \frac{\sin \alpha}{\sin (\alpha + \phi)} \geq \sin \alpha > 0. \tag{7.5}$$

As for the second part, since θ is not a point of $\mathcal{E}$, there exists a $\nu = \nu(\theta)$ such that $\theta_1(\nu) < \theta < \theta_2(\nu)$, where $P_1(\nu) = e^{i\theta_1}$, $P_2(\nu) = e^{i\theta_2}$. We construct

angles OQR and RP_1S each equal to α, and then triangles OQR and RSP_1 are similar. We have, writing $\psi =$ angle $P_1SR =$ angle ROQ,

$$\frac{P_1S}{\sin(\alpha+\psi)} = \frac{P_1R}{\sin\psi} = \frac{1-OR}{\sin\psi}, \tag{7.6}$$

and

$$OR = \frac{\sin\alpha}{\sin(\alpha+\psi)}. \tag{7.7}$$

Substituting from (7.7) into (7.6) and simplifying, we obtain

$$P_1S = \cos\alpha - \sin\alpha\cdot\tan\psi/2, \tag{7.8}$$

and it follows that, if $\psi < \psi_0(\alpha)$, then

$$P_1S \geq K_2(\alpha) > 0. \tag{7.9}$$

Hence if $\psi < \psi_0(\alpha)$, part (ii) of Lemma 1 is true. But if $\psi \geq \psi_0(\alpha)$, (ii) is obviously true, and this completes the proof of the lemma.

By (i) of Lemma 1, $z = 0$ is an interior point of D^*, and there exists a function $h(z_1)$ which maps the unit circle $|z_1| < 1$ on D^* with $h(0) = 0$. We write $F(z_1) = f\{h(z_1)\}$, $z_1 = \rho e^{i\phi}$. It is, then, sufficient to prove that

$$\int_0^1 S_F(r)\,dr = \int_0^1 \left\{ \frac{1}{\pi}\int_0^r \int_{-\pi}^{\pi} \frac{|F'|^2}{(1+|F|^2)^2}\,\rho d\rho d\phi \right\} dr < \infty. \tag{7.10}$$

For then $T_F(r) = O(1)$ by (1.7), and so, by (N), for almost all ϕ, F tends to a limit as $z_1 \to e^{i\phi}$ in any $\Gamma_\gamma(\phi)$ for which $\gamma < \pi/2$. Also $h(z_1)$, a bounded function, tends to a limit for almost all ϕ as $z_1 \to e^{i\phi}$ in $\Gamma_\gamma(\phi)$, $\gamma < \pi/2$. Since the boundary of D^* is rectifiable, it has a tangent almost everywhere. Hence almost everywhere the angle between the right-hand and left-hand tangents is π, and it follows at once that $\lim f(z)$ exists in every $\Gamma_\beta(\theta)$, $\beta < \pi/2$, almost everywhere in $\mathcal{E}$.

Let $t(r)$ be the transform of $|z| < r$ by h, and let $D^*(r) = D^* \cap C_r$ (C_r being the circle with centre at $z = 0$ and radius r). Then since $|h| \leq 1$, $h(0) = 0$,

$$t(r) \subset D^*(r)$$

by Schwarz's lemma, and so

$$\begin{aligned} S_F(r) &= \frac{1}{\pi}\int_0^r \int_{-\pi}^{\pi} \frac{|F'|^2}{(1+|F|^2)^2}\,\rho d\rho d\phi \\ &= \frac{1}{\pi}\iint_{t(r)} \frac{|f'|^2}{(1+|f|^2)^2}\,\rho d\rho d\theta \leq \iint_{D^*(r)} \frac{|f'|^2}{(1+|f|^2)^2}\,\rho d\rho d\theta. \end{aligned} \tag{7.11}$$

Next, writing

$$\psi(\rho,\theta)=\begin{cases}\dfrac{|f'|^2}{(1+|f|^2)^2}, & \text{if } (\rho,\theta)\subset D^* \\ 0, & \text{otherwise,}\end{cases}$$

we have, for suitable $K_3=K_3(\alpha)$,

$$(7.12)\quad \int_0^1\int_{\theta-K_3\cdot(1-\rho)}^{\theta+K_3\cdot(1-\rho)}\psi(\rho,\phi)\rho d\rho d\phi\leq\iint_{\Gamma_\alpha(\theta)}\psi(\rho,\phi)\rho d\rho d\phi=\sum_{k=1}^{4}\iint_{\Gamma_k}\leq 2K_1+K_4(f,$$

if $\theta\notin\mathcal{E}$, by (ii) of Lemma 1 and (7.1); and (7.12) is true *a fortiori* if $\theta\subset\mathcal{E}$. Hence by (7.11) and (7.12) [9]

$$(7.13)\quad \int_0^1 S_F(r)\,dr\leq\int_0^1\left\{\iint_{D^*(r)}\frac{|f'|^2}{(1+|f|^2)^2}\rho d\rho d\phi\right\}dr$$

$$=\int_0^1\int_{-\pi}^{\pi}(1-\rho)\cdot\psi(\rho,\phi)\rho d\rho d\phi$$

$$=\frac{1}{2K_3}\int_{-\pi}^{\pi}d\theta\left\{\int_0^1\int_{\theta-K_3\cdot(1-\rho)}^{\theta+K_3\cdot(1-\rho)}\psi(\rho,\phi)\rho d\rho d\phi\right\}\leq\frac{\pi}{K_3}\{2K_1+K_4\}<\infty.$$

This proves (7.10), and completes the proof of Theorem 5.

STANFORD UNIVERSITY, CALIFORNIA.

REFERENCES

1. Ahlfors, L., "Beitrage zur Theorie der meromorphen Funktionen," 7. *Congr. Math. Scand.*, Oslo (1929).
2. Ahlfors, L., "Über eine Methode in der Theorie der meromorphen Funktionen," *Soc. sci. fenn. Comment. Phys.-math.* 8, Nr. 10 (1935).
3. Shimizu, T., "On the theory of meromorphic functions," *Jap. Jour. Math.*, vol. 6 (1929), pp. 119-171.
4. Spencer, D. C., "On finitely mean valent functions," *Proc. London Math. Soc.*, vol. 47 (1941), pp. 201-211.
5. Spencer, D. C., "On finitely mean valent functions II," *Trans. Amer. Math. Soc.*, vol. 48 (1940), pp. 418-435.
6. Spencer, D. C., "Note on some function-theoretic indentities," *Jour. London Math. Soc.*, vol. 15 (1940), pp. 84-86.
7. Spencer, D. C., "On mean one-valent functions," *Annals of Math.*, vol. 42 (1941), pp. 614-633.

[9] These final steps in the proof, as well as the use of the domain D^*, were suggested by some remarks of Professor J. E. Littlewood in a letter to Professor Zygmund.

ON FUNCTIONS OF CLASS 1 HAVING THE PROPERTY OF DARBOUX.*

By ISAIAH MAXIMOFF.

1. *Definition* 1. A point x_0 of a perfect set P is called a point of second species of P if it is a both-sided limit point of P. Let E be a linear B-measurable set.

Definition 2. We shall say that a point x_0 of E is a point of second species of E in the segment $[a, b]$ if there exists a perfect set $P(x_0)$ having the following properties:

(i) $P(x_0)$ contains x_0 and is contained in $[a, b]E$;

(ii) if $a < x_0 < b$, then x_0 is a point of second species of $P(x_0)$.

Let $f(x)$ be a finite function of class 1 on $[a, b]$.

Definition 3. We shall say that a perfect set $P(x_0)$ is a *perfect road* for $f(x)$ in $[a, b]$ at the point x_0 if

(i) x_0 is a point of second species of $P(x_0)$ in $[a, b]$;

(ii) $f(x)$ is continuous at the point x_0 relative to $P(x_0)$.

The author has established the following theorem:

If a finite function $f(x)$ of Baire's classification has the property of Darboux in $[a, b]$, then there exists for every point x_0 of $[a, b]$ a perfect road $Q(x_0)$ for $f(x)$ in $[a, b]$.[1]

Definition 4. A finite function $f(x)$ will be called *partially continuous* at the point x_0 of the segment $[a, b]$ if there exists at x_0 a perfect road $P(x_0)$ for $f(x)$ in $[a, b]$.

The author has also proved the following theorem:

A necessary and sufficient condition that a finite function $f(x)$ of class 1 have the property of Darboux in $[a, b]$ is that $f(x)$ be partially continuous at each point of the segment.[2] This theorem will be called the *first theorem on*

* Received May 3, 1939; Revised August 14, 1940.

[1] "Sur les fonctions ayant la propriété de Darboux," *Prace Mat.-Fiz.*, vol. 43 (1935), pp. 241-242, 258-259.

[2] *Loc. cit.*[1], p. 260.

functions of class 1 having the property of Darboux. There is an analogy between this theorem and a well-known theorem of Baire on functions of class 1: *La condition nécessaire et suffisante pour qu'une fonction quelconque finie ou infinie soit limite de fonctions continues, est qu'elle soit ponctuellement discontinue sur tout ensemble parfait.*[3]

Furthermore there is a well-known theorem of Lebesgue on functions of class 1:

La condition nécessaire et suffisante pour que f, supposée discontinue, soit de classe 1 *sur P, est que les ensembles de points de P,* $E(f > A)$, $E(f < A)$, *soient des sommes d'ensembles fermés quelque soit A.*[4]

2. We thus have two theorems on functions of class 1, a) Baire's theorem and b) Lebesgue's theorem. The author has obtained the analogue of Baire's theorem for functions of class 1 having the property of Darboux; it is of interest to obtain the second theorem on these functions analogous to Lebesgue's theorem. In order to formulate and to prove this theorem, we consider any two rational numbers w, v $(w < v)$ and the set E_v^w of all points of the segment $[a, b]$ such that $w < f(x) < v$.

LEMMA 1. *If* $f(x)$ *is a finite function of class* 1 *having the property of Darboux in* $[a, b]$ *and if* x_0 *is a point of the set* E_v^w, *then there exists a non-dense perfect set* $P(x_0)$ *such that*

(i) $P(x_0)$ *is contained in the set* E_v^w;
(ii) x_0 *is a point of second species of* $P(x_0)$ *in* $[a, b]$;
(iii) $f(x)$ *is continuous at the point* x_0 *relative to* $P(x_0)$.

It is obvious that $P(x_0)$ is a perfect non-dense road for $f(x)$ in $[a, b]$ at the point x_0.

Proof. Since $f(x)$ is a finite function of class 1 having the property of Darboux in $[a, b]$ we can find a perfect road $Q(x_0)$ for $f(x)$ in $[a, b]$ at the point x_0. Let x_0 be any point of E_v^w. This means that $w < f(x_0) < v$. The function $f(x)$ is continuous at the point x_0 relative to $Q(x_0)$ and so we can find a segment $s = [x_0 - \epsilon, x_0 + \epsilon]\,[a, b]$, where $\epsilon > 0$, such that $sQ(x_0) \subset E_v^w$. Let $R(x_0)$ denote the largest perfect set contained in $sQ(x_0)$. It is readily seen that $R(x_0)$ is a perfect road at the point x_0 for $f(x)$ in $[a, b]$ contained in E_v^w. Evidently, there exists a non-dense perfect set $P(x_0)$ such that

---

[3] R. Baire, *Leçons sur les fonctions discontinues* (1930), p. 124.

[4] H. Lebesgue, "Une propriété caractéristique des fonctions de classe *un*," *Bull. de la Soc. Math. de France*, vol. 32 (1904), pp. 229-242.

(i) $P(x_0)$ is contained in $R(x_0)$;

(ii) x_0 is a point of second species of $P(x_0)$ in $[a, b]$.

This set $P(x_0)$ satisfies all the conditions of our lemma. $P(x_0)$ will be called a perfect non-dense road for $f(x)$ at the point x_0.

If η is any positive number, let $E[f(x_0), \eta]$ denote the set of all points x of $[a, b]$ such that $| f(x) - f(x_0) | < \eta$.

LEMMA 2. *A necessary and sufficient condition that the function $f(x)$ be partially continuous at the point x_0 of the segment $[a, b]$ is that x_0 be a point of second species of $E[f(x_0), \eta]$ in $[a, b]$ for every positive η.*

Proof. x_0 is a point of second species of $E_n = E[f(x_0), 1/n]$ in $[a, b]$ and so there exists a perfect set P_n such that

(i) x_0 is a point of second species of P_n in $[a, b]$;

(ii) P_n is contained in E_n.

Now we form the sets

$$H'_n = (x_0 - 1/n, x_0)P_n, \qquad H''_n = (x_0, x + 1/n)P_n$$

and the sets G'_n, G''_n in such a manner that

(i) if the set $H'_n - H'_{n+1}\{H''_n - H''_{n+1}\}$ is uncountable, then $G'_n\{G''_n\}$ is a perfect set contained in $H'_n - H'_{n+1}\{H''_n - H''_{n+1}\}$;

(ii) if the set $H_n' - H_n'_{+1}\{H''_n - H''_{n+1}\}$ is at most countable, then $G'_n\{G''_n\}$ is null.

If $a < x_0 < b$, then the set $x_0 + \sum_{n=1}^{n=\infty}(G'_n + G''_n)$ is a perfect road for $f(x)$ in $[a, b]$ at the point x_0. If $x_0 = a\{x_0 = b\}$, then the set $x_0 + \sum_{n=1}^{n=\infty} G''_n$ $\{x_0 + \sum_{n=1}^{n=\infty} G'_n\}$ is the perfect road for $f(x)$ in $[a, b]$ at the point $x_0 = a\{x_0 = b\}$. From this we conclude at once that $f(x)$ is partially continuous at each point x_0 of $[a, b]$.

We next assume that $f(x)$ is partially continuous at a point x_0 of $[a, b]$. This means that there exists a perfect road $P(x_0)$ for $f(x)$ in $[a, b]$ at the point x_0. Let η be any positive number. We form the set $P_n(x_0) = P(x_0) \cdot E[f(x_0), \eta]$. The function $f(x)$ is continuous at the point x_0 over $P(x_0)$ so that we can find a segment $s = [x_0 - \epsilon, x_0 + \epsilon][a, b]$, where $\epsilon > 0$, such that $sP(x_0) = s \cdot P_n(x_0)$. Since x_0 is a point of second species of $P(x_0)$ in $[a, b]$, x_0 is a point of second species of $sP_\eta(x_0)$ in $[a, b]$. But $sP_\eta(x_0)$ is contained in $E[f(x_0), \eta]$, so that x_0 is a point of second species of $E[f(x_0), \eta]$ in $[a, b]$.

3. Let (y) $y_1, y_2, y_3, \cdots$ be the sequence of all rational numbers and consider any system of perfect sets

$$(H) \qquad H_{y_\rho}^{y_r}[r+\rho+s] \qquad \begin{pmatrix} s=0,1,2,3,\cdots \\ r=1,2,3,4,\cdots \\ \rho=1,2,3,4,\cdots \end{pmatrix}$$

where $y_r < y_\rho$. We shall say that this system H possesses

a) the property $(E = lH)$ if

$$(1) \qquad E_{y_\rho}^{y_r} = \lim_{s=\infty} H_{y_\rho}^{y_r}[r+\rho+s];$$

b) the property $(H \subset E)$ if

$$(2) \qquad H_{y_\rho}^{y_r}[r+\rho+s] \subset E_{y_\rho}^{y_r};$$

c) the property $(H \underset{s}{\subset} H)$ if

$$(3) \qquad H_{y_\rho}^{y_r}[r+\rho+s] \subset H_{y_\rho}^{y_r}[r+\rho+s+1];$$

d) the property $(H \underset{2}{\subset} H)$ if every point of $H_{y_\rho}^{y_r}[M+s]$ where M is the greater of the integers $r+\rho$, $t+\tau$, is a point of second species whenever $y_t < y_r < y_\rho < y_\tau$.

We now state the second theorem on functions of class 1 having the property of Darboux.

THEOREM 1. *A necessary and sufficient condition that a finite function $f(x)$ of class* 1 *have the property of Darboux is that there exists a system of perfect sets*

$$(P) \qquad P_{y_\rho}^{y_r}[r+\rho+s] \qquad \begin{pmatrix} s=0,1,2,3,\cdots \\ r=1,2,3,4,\cdots \\ \rho=1,2,3,4,\cdots \end{pmatrix}$$

having the properties:

$$\text{a) } (E = lP), \quad \text{b) } (P \subset E), \quad \text{c) } (P \underset{s}{\subset} P), \quad \text{d) } (P \underset{2}{\subset} P).$$

Proof. Let $f(x)$ be a finite function of class 1, then $E_{y_\rho}^{y_r}$ is the sum of perfect sets Q_n $(n = 1,2,3,4,\cdots)$ and of a countable set of points x_n $(n = 1,2,3,4,\cdots)$. On the other hand, $f(x)$ possesses the property of Darboux in $[a,b]$ so that we can find by Lemma 1 a perfect road $P(x_n)$ at the point x_n for $f(x)$ contained in $E_{y_\rho}^{y_r}$. It is clear that $E_{y_\rho}^{y_r}$ is the sum of perfect sets Q_m and $P(x_n)$ $\{m = 1,2,3,4,\cdots;\ n = 1,2,3,4,\cdots\}$.

We may thus write

$$(4) \qquad E_{y_\rho}^{y_r} = \sum_{k=1}^{\infty} A_{y_\rho}^{y_r}[r+\rho+k]$$

where $A^{\nu_r}_{\nu_\rho}[r+\rho+k]$ are the perfect sets. Hence, we obtain

$$E^{\nu_r}_{\nu_\rho} = \lim_{s=\infty} H^{\nu_r}_{\nu_\rho}[r+\rho+s] \tag{5}$$

where $H^{\nu_r}_{\nu_\rho}[r+\rho+s]$ are perfect sets defined by the relations

$$\begin{cases} H^{r+s}_{\nu_r} = A^1{}_{\nu_r} + A^2{}_{\nu_r} + \cdots + A^{r+s}_{\nu_r} \\ H^{\nu_r}_{r+s} = A_1{}^{\nu_r} + A_2{}^{\nu_r} + \cdots + A^{\nu_r}_{r+s}\,. \end{cases} \tag{6}$$

It is obvious that the system H of these perfect sets $H^{\nu_r}_{\nu_\rho}[r+\rho+s]$ possesses the properties $(E = lH)$ and $(H \subset E)$.

Now we consider any system of perfect sets

$$(H) \qquad H^{\nu_r}_{\nu_\rho}[r+\rho+s] \qquad \begin{pmatrix} s = 0,1,2,3,4,\cdots \\ r = 1,2,3,4,5,\cdots \\ \rho = 1,2,3,4,5,\cdots \end{pmatrix}$$

which posseses the properties $(E = lH)$ and $(K \subset E)$.

Let P be an operation which transforms the system H of sets

$$(H) \qquad H^{\nu_r}_{\nu_\rho}[r+\rho+s] \qquad \begin{pmatrix} s = 0,1,2,3,\cdots \\ r = 1,2,3,4,\cdots \\ \rho = 1,2,3,4,\cdots \end{pmatrix}$$

into a system PH of perfect sets

$$(PH) \qquad PH^{\nu_r}_{\nu_\rho}[r+\rho+s].$$

Here the system PH is the result of the operation P on the system H. We denote by P_2P_1H the system of perfects sets which is the result of an operation P_2 on the system P_1H. It is obvious that the system P_2P_1H may be considered as the result of the operation P_2P_1 on the system H.

We shall now find an operation P which transforms the system H into the system P of perfect sets

$$(P) \qquad P^{\nu_r}_{\nu_\rho}[r+\rho+s] \qquad \begin{pmatrix} s = 0,1,2,3,\cdots \\ r = 1,2,3,4,\cdots \\ \rho = 1,2,3,4,\cdots \end{pmatrix}$$

where

$$P^{\nu_r}_{\nu_\rho}[r+\rho+s] = PH^{\nu_r}_{\nu_\rho}[r+\rho+s] \tag{7}$$

satisfying all the conditions of our theorem. Formula (7) may be written in the form $P = PH$.

We define the operation P by the formula

$$P = \cdots \mathfrak{P}_n \cdots \mathfrak{P}_5 \cdot \mathfrak{P}_4 \cdot \mathfrak{P}_3 \cdot \mathfrak{P}_2, \tag{8}$$

where $\mathfrak{P}_n$ is an operation such that if the system H possesses properties $(E = lH)$ and $(H \subset E)$, then the system $\mathfrak{P}_n H$ of perfect sets

$$(\mathfrak{P}_n H) \qquad \mathfrak{P}_n H^{y_r}_{y_\rho}[r+\rho+s] \qquad \begin{pmatrix} s = 0, 1, 2, 3, \cdots \\ r = 1, 2, 3, 4, \cdots \\ \rho = 1, 2, 3, 4, \cdots \end{pmatrix}$$

posseses

a_1) the property $(E = l\mathfrak{P}_n H)$, i. e. $E^{y_r}_{y_\rho} = \lim_{s=\infty} \mathfrak{P}_n H^{y_r}_{y_\rho}[r+\rho+s]$;

b_1) the property $(H \subset \mathfrak{P}_n H \subset E)$, i. e.

$$H^{y_r}_{y_\rho}[r+\rho+s] \subset \mathfrak{P}_n H^{y_r}_{y_\rho}[r+\rho+s] \subset E^{y_r}_{y_\rho};$$

c_1) the property $(\mathfrak{P}_n H \underset{s}{\subset} \mathfrak{P}_n H)_k$ for all integers $k \leqq n$, i. e.

$$\mathfrak{P}_n H^{y_r}_{y_\rho}[r+\rho+s] \subset \mathfrak{P}_n H^{y_r}_{y_\rho}[r+\rho+s+1]$$

whenever $r+\rho+s+1 \leqq 2k \leqq 2n$;

d_1) the property $(\mathfrak{P}_n H \underset{2}{\subset} \mathfrak{P}_n H)_k$ for all integers $k \leqq n$, i. e. every point of the set $\mathfrak{P}_n H^{y_r}_{y_\rho}[2k]$ is a point of second species of the set $\mathfrak{P}_n H^{y_t}_{y_\tau}[2k]$ whenever $k \leqq n$,

$$y_t < y_r < y_\rho < y_\tau, \quad t \leqq k, \quad r \leqq k, \quad \rho \leqq k, \quad \tau \leqq k;$$

e_1) if the system of sets

$$H^{y_r}_{y_\rho}[r+\rho+s] \qquad \begin{pmatrix} s = 0, 1, 2, 3, \cdots \\ r = 1, 2, 3, 4, \cdots \\ \rho = 1, 2, 3, 4, \cdots \end{pmatrix}$$

possesses properties $(E = lH)$, $(H \subset E)$, $H \underset{s}{\subset} H)_k$ and $(H \underset{2}{\subset} H)_k$ for $k \leqq n$, then for all $k \leqq n$ we have $(\mathfrak{P}_n H = H)_k$, i. e. $\mathfrak{P}_n H^{y_r}_{y_\rho}[2k] = H^{y_r}_{y_\rho}[2k]$ whenever $k \leqq n$, $r \leqq k$, $\rho \leqq k$.

We define the operation $\mathfrak{P}_n$ by the formula

$$(9) \qquad \mathfrak{P}_n = R_n S_n T_n$$

as the combination of certain operations T_n, S_n. R_n. We define the operation T_n by the formulas

$$(10) \qquad \begin{cases} T_n H^{y_r}_{y_\rho}[r+\rho+s] = H^{y_r}_{y_\rho}[r+\rho] + H^{y_r}_{y_\rho}[r+\rho+1] + \cdots \\ \qquad + H^{y_r}_{y_\rho}[r+\rho+s] \text{ for all } r+\rho+s \leqq 2n; \end{cases}$$

$$(11) \qquad \begin{cases} T_n H^{y_r}_{y_\rho}[r+\rho+s] = H^{y_r}_{y_\rho}[r+\rho+s] \\ \qquad \text{for all } r+\rho+s > 2n. \end{cases}$$

It is easily seen tha tthe operation T_n possesses the properties

a_2) $(E = lT_nH)$; b_2) $(H \subset T_nH \subset E)$; c_2) $(T_nH \underset{s}{\subset} T_nH)_k$ for all $k \leqq n$; d_2) if for all $k \leqq n$, $(H \underset{s}{\subset} H)_k$, then $(T_nH = H)_k$ for all $k \leqq n$, i. e.,

$$T_nH^{y_r}_{y_\rho}[r+\rho+s] = H^{y_r}_{y_\rho}[r+\rho+s] \text{ for all } r+\rho+s \leqq 2k \leqq 2n.$$

We denote the set $T_nH^{y_r}_{y_\rho}[r+\rho+s]$ by $T^{y_r}_{y_\rho}[r+\rho+s]$, i. e. $TH = T$.

We now define the operation S_n by the formulas

$$(12) \quad \begin{cases} S_nT^{y_t}_{y_\rho}[2n] = \sum\limits_{r\rho} T^{y_r}_{y_\rho}[2n] \text{ whenever} \\ \qquad r \leqq n,\ t \leqq n,\ \rho \leqq n,\ \tau \leqq n;\ y_t \leqq y_r \leqq y_\rho \leqq y_\tau; \end{cases}$$

$$(13) \quad \begin{cases} S_nT^{y_t}_{y_\rho}[t+\tau+s] = T^{y_t}_{y_\rho}[t+\tau+s] \\ \qquad \text{in all other cases.} \end{cases}$$

Here $\sum\limits_{r\rho}$ extends over all integers r, ρ for which $y_t \leqq y_r \leqq y_\rho \leqq y_\tau$, $r \leqq n$, $t \leqq n$, $\tau \leqq n$. We readily verify that the operation S_n posseses the properties

a_3) $(E = lS_nT)$; b_3) $(T \subset S_nT \subset E)$; c_3) if $(T \underset{s}{\subset} T)_k$ for all $k \leqq n$, then $(S_nT \subset S_nT)_k$ for all $k \leqq n$; d_3) if for all $k \leqq n$ we have $(T \underset{y}{\subset} T)_k$, i. e.

$$T^{y_r}_{y_\rho}[2k] \subset T^{y_t}_{y_\tau}[2k] \text{ whenever}$$

$$y_t < y_r < y_\rho < y_\tau, \quad t \leqq k, \quad r \leqq k, \quad \rho \leqq k, \quad \tau \leqq k \leqq n,$$

then for all $k \leqq n$ we have $(S_nT = T_k)$, i. e.

$$S_nT^{y_r}_{y_\rho}[2k] = T^{y_r}_{y_\rho}[2k].$$

We define the operation R_n in the following manner. We first denote the operation S_nT_n by V_n and define the system N of sets

$$(N) \qquad N^{y_r}_{y_\rho}[r+\rho+s] \qquad \begin{pmatrix} s = 0, 1, 2, 3, \cdots \\ r = 1, 2, 3, 4, \cdots \\ \rho = 1, 2, 3, 4, \cdots \end{pmatrix}$$

by the formula

$$(14) \qquad N^{y_r}_{y_\rho}[r+\rho+s] = V_nH^{y_r}_{y_\rho}[r+\rho+s].$$

Evidently, V_n posseses the properties

a_4) $(E = lV_nH)$; b_4) $(H \subset V_nH \subset E)$; c_4) $(V_nH \underset{s}{\subset} V_nH)_k$ for all $k \leqq n$; d_4) $(V_nH \underset{y}{\subset} V_nH)_k$ for all $k \leqq n$, because $V_nH = S_nT$, and therefore

the properties a_4), b_4), c_4), d_4) coincide respectively with the properties a_3), b_3), c_3), d_3).

Now we suppose that $y_t < y_r < y_\rho < y_\tau$, $t \leqq n$, $r \leqq n$, $\rho \leqq n$, $\tau \leqq n$ and that $X^{y_r}_{y_\rho}[2n]$, $X^{y_t}_{y_\tau}[2n]$ are any perfect sets such that

$$X^{y_r}_{y_\rho}[2n] \subset X^{y_t}_{y_\tau}[2n] \subset E^{y_t}_{y_\tau}.$$

The set of all points $x^{rt}_{\rho\tau}$ which belong to the set $X^{y_r}_{y_\rho}[2n]$ and which are not points of second species of the set $X^{y_t}_{y_\tau}[2n]$ in $[a, b]$ is countable. It is clear that $x^{rt}_{\rho\tau}$ is the extremity of a contiguous interval (a, b) of the set $X^{y_t}_{y_\tau}[2n]$. Let $P(x^{rt}_{\rho\tau})$ be a non-dense perfect road at the point $x^{rt}_{\rho\tau}$ contained in $X^{y_t}_{y_\tau}[2n]$. We now adjoin to the set $X^{y_t}_{y_\tau}[2n]$ the set $[a, b]X^{y_t}_{y_\tau}[2n]$ or the set $[a, b]$. $P(x^{rt}_{\rho\tau})$ according as the point $x^{rt}_{\rho\tau}$ is or is not a point of second species of the set $X^{y_t}_{y_\tau}[2n]$. Denote by $A(x^{rt}_{\rho\tau})$ the result of this adjunction and consider the composed operation $A^{rt}_{\rho\tau}$ which consists of the $A(x^{rt}_{\rho\tau})$ for all points $x^{rt}_{\rho\tau}$.

Let $y_{a_1}, y_{a_2}, \cdots y_{a_n}$ be an increasing sequence of the numbers $y_1, y_2, \cdots y_n$. We shall define the operation R_n by the formula

$$R_n = A^{a_{n-2}a_{n-3}}_{a_{n-1}a_n} \cdot S_n \cdot \cdots A^{a_4 \cdot a_3}_{a_5 \cdot a_6} \cdot S_n \cdot A^{a_3a_2}_{a_4a_5} \cdot S_n \cdot A^{a_2a_1}_{a_3a_4} \cdot S_n.$$

Evidently, the operation R_n possesses the properties

a_5) $(E = lR_nN)$; b_5) $(N \subset R_nN \subset E)$; c_5) $(R_n \underset{s}{\subset} R_nN)_k$ for all $k \leqq n$, because in virtue of c_4) we have $(V_nH \underset{s}{\subset} V_nH)_k$ or $(N \underset{s}{\subset} N)_k$ for all $k \leqq n$ and the operation R_n changes only the sets $N^{y_r}_{y_\rho}[r + \rho + s]$ of order n, $r + \rho + s = 2n$, in such manner that

$$N^{y_r}_{y_\rho}[2n] \subset R_nN^{y_r}_{y_\rho}[2n];$$

d_5) $(R_nN \underset{2}{\subset} R_nN)_k$ for all $k \leqq n$; e_5) since $(E = lN)$, $(N \subset E)$, $(N \underset{s}{\subset} N)_s$ and $(N \underset{2}{\subset} N)_k$ for all $k \leqq n$, then $(R_nN = N)_k$ for all $k \leqq n$. It is easily shown that the operation $\mathfrak{P}_n = R_nS_nT_n$ possesses all the properties a_1), b_1), c_1), d_1), e_1). For this purpose it is sufficient in a_5), b_5), c_5), d_5), e_5) to substitute R_nN instead of $\mathfrak{P}_nH$ because $R_nN = \mathfrak{P}_nH$.

It is obvious that the operation

$$P = \cdots \mathfrak{P}_n \cdots \mathfrak{P}_5 \cdot \mathfrak{P}_4 \cdot \mathfrak{P}_3 \cdot \mathfrak{P}_2$$

furnishes a system of perfect sets

$$(P) \qquad P^{y_r}_{y_\rho}[r + \rho + s] \qquad \begin{pmatrix} s = 0, 1, 2, 3, \cdots \\ r = 1, 2, 3, 4, \cdots \\ \rho = 1, 2, 3, 4, \cdots \end{pmatrix}$$

where

$$P^{y_r}_{y_\rho}[r+\rho+s] = PH^{y_r}_{y_\rho}[r+\rho+s]$$

which satisfies all the conditions of our theorem. Besides, the operation P possesses the following property: if $(E = lH)$, $(H \subset E)$, $(H \underset{s}{\subset} H)_k$ and $(H \underset{2}{\subset} H)_k$ for all $k \leqq n$, then $PH^{y_r}_{y_\rho}[r+\rho+s] = H^{y_r}_{y_\rho}[r+\rho+s]$. Thus, the conditions of our theorem are necessary. To prove that these conditions are sufficient, let x_0 be any point of the segment $[a, b]$ and let $y_0 = f(x_0)$. Then, if η is a sufficiently small positive number we have integers t, τ, r, ρ such that

$$y_0 - \eta < y_t < y_r < y_0 < y_\rho < y_\tau < y_0 + \eta.$$

The point x_0 belongs to the set $E^{y_r}_{y_\rho} \cdot E^{y_t}_{y_\tau}$. Consequently, x_0 belongs to the set

$$P^{y_r}_{y_\rho}[2n] \cdot P^{rt}_{y_\tau}[2n]$$

for sufficiently large n if $r \leqq n$, $\rho \leqq n$, $t \leqq n$, $\tau \leqq n$. Therefore x_0 is a point of second species of $P^{y_t}_{y_\tau}[2n]$. But if $E[f(x_0), \eta]$ is the set of all points x of $[a, b]$ satisfying the condition $|f(x) - f(x)| < \eta$, we have $P^{rt}_{y_\tau}[2n] \subset E[f(x_0), \eta]$ so that x_0 is a point of second species of $E[f(x_0), \eta]$ in $[a, b]$. Consequently, in virtue of Lemma 2, the function $f(x)$ is partially continuous at each point x_0 of $[a, b]$. But $f(x)$ is a finite function of class 1, therefore from the first theorem on functions of class 1 having the property of Darboux it follows that $f(x)$ possesses the property of Darboux.

Our theorem is thus proved.

4. We now turn to the functional equation

$$g(x) = f[\phi(x)]$$

where $g(x)$ and $f(y)$ are given functions, and $\phi(x)$ is sought.[5] We consider the case where $\phi(x)$ possesses the property of Darboux in the segment $[a, b]$. Consider the sequence of all rational numbers (y) $y_1, y_2, y_3, \cdots$. Let $E^{y_r}_{y_\rho}$ be the set of all points (x, y) such that

$$g(x) = f(y), \qquad a \leqq x \leqq b, \qquad y_r < y < y_\rho.$$

Denote by PE the orthogonal projection of a set E on the x axis. Suppose now that M is the greater of the integers $\rho + r$, $\tau + t$ and assume $y_t < y_r < y_\rho < y_\tau$.

---

[5] S. Braun, "Sur l'equation fonctionnelle $g(x) = f[\phi(x)]$," *Fundamenta Mathematicae*, vol. 28 (1937), pp. 294-302; W. Sierpinski, "Sur les fonctions dépendantes," *Fundamenta Mathematicae*, vol. 28 (1936), pp. 66-70.

THEOREM 2. *A necessary condition that there exist a finite function* $\phi(x)$ *of class* 1 *having the property of Darboux and satisfying the equation* $g(x) = f[\phi(x)]$ *is that there exists a system* P *of perfect sets* (P) $P^{y_r}_{y_\rho}[r+\rho+s]$ *such that*

(i) $$\lim_{s=\infty} P^{y_r}_{y_\rho}[r+\rho+s] \subset PE^{y_r}_{y_\rho};$$
$$P^{y_r}_{y_\rho}[r+\rho\, s] \subset P^{y_r}_{y_\rho}[r+\rho+s+1] \subset PE^{y_r}_{y_\rho};$$

(ii) *every point of* $P^{y_r}_{y_\rho}[M+S]$ *is a point of second species of the set* $P^{r_t}_{y_\tau}[M+S]$;

(iii) $$\lim_{s=\infty} \sum_{r\rho} P^{y_r}_{y_\rho}[r+\rho+s] = [a,b]$$

Proof. We assume that a finite function of class 1 having the property of Darboux satisfies the functional equation $g(x) = f[\phi(x)]$. It is clear that every point $[x, \phi(x)]$ belongs to the set $E^{y_r}_{y_\rho}$. Since $\phi(x)$ possesses the property of Darboux, then there exists a system P of perfect sets

(P) $$P^{y_r}_{y_\rho}[r+\rho+s], \qquad y_r < y_\rho,$$

satisfying the second condition of Theorem 1. Besides, it is easily seen that

$$\lim_{s=\infty} P^{y_r}_{y_\rho}[r+\rho+s] \subset PE^{y_r}_{y_\rho},$$

$$P^{y_r}_{y_\rho}[r+\rho+s] \subset P^{y_r}_{y_\rho}[r+\rho+s+1] \subset PE^{y_r}_{y_\rho},$$

$$\lim_{s=\infty} \sum_{r\rho} P^{y_r}_{y_\rho}[r+\rho+s] = [a,b].$$

Thus, our theorem is proved.

TCHEBOKSARY, U. S. S. R.

THE CALCULUS OF NON-CONTRADICTION.*

By Glenn Edward Dexter.

In this paper an especially simple calculus of propositions, requiring only one assumption, is presented. The principal symbols used are the letters $p, q, r, \cdots$ and a horizontal line, "——." The line forms a *parenthesis*—thus $\overline{p \quad q}$ is a parenthesis in which p and q occur; $\bar{p}$ is a parenthesis in which only p occurs; $\overline{p \quad q \quad \bar{r}}$ is a parenthesis in which p, q and $\bar{r}$ occur. It should be noted that r does NOT occur in $\overline{p \quad q \quad \bar{r}}$. The accompanying table correlates the new notation with ordinary language and with the notation of *Principia Mathematica*.[1] We see that "not (not p)" becomes $\bar{\bar{p}}$; "if p and q, then r" becomes $\overline{\overline{p \quad q} \quad \bar{r}}$; and so on.

Table

not p	$\sim p$	$\bar{p}$
if p, then q	$p \supset q$	$\overline{p \quad \bar{q}}$
p or q	$p \vee q$	$\overline{\bar{p} \quad \bar{q}}$
p or q or r	$p \vee q \vee r$	$\overline{\bar{p} \quad \bar{q} \quad \bar{r}}$
p and q	$p \cdot q$	$\overline{\overline{p \quad q}}$
p and q and r	$p \cdot q \cdot r$	$\overline{\overline{p \quad q \quad r}}$

The word *proposition* will denote, indifferently, a letter or a parenthesis. To say that we *have* a certain proposition will mean that that proposition can be written without being covered by a line. Thus if we write

$$p \quad \bar{q} \quad \overline{\bar{r} \quad \bar{s}}$$

* Received August 29, 1941; Revised February 16, 1942.

[1] A. N. Whitehead and B. Russell, *Principia Mathematica*, vol. I, 1910. Cambridge University Press.

we evidently have three propositions—p, $\bar{q}$ and $\overline{\bar{r}\ \bar{s}}$. To say that we have a *contradiction* will mean that we have a parenthesis, P, and have also every proposition which occurs in P. Thus we have a contradiction in

$$p \quad \overline{p\ q\ \bar{r}} \quad q \quad r,$$

or in

$$\overline{\overline{\bar{q}\ \bar{r}}} \quad p \quad \overline{\overline{\bar{q}\ \bar{r}}} \quad p.$$

The working of the calculus will depend upon one assumption, as follows: *If P is a parenthesis such that if we had P we would have a contradiction, then we have every proposition which occurs in P.*

A calculation may or may not start from an hypothesis. In either case, the object is to prove that we have one or more propositions designated *conclusions*. (Where there is no hypothesis, as in Example 8, the conclusions will be universally valid propositions.) This can be done, in accordance with the assumption, by showing that there is a parenthesis, P, in which the conclusions occur, and such that if we had P we would have a contradiction. This parenthesis will be called the *trial parenthesis*. Sometimes, as in Example 2, it will necessarily contain propositions other than the conclusions.

Example 1. Given $\sim(\sim p)$; to prove that we have p.

$$\bar{\bar{p}} \quad \bar{p}$$

(1) (2)

We have (1) by hypothesis. Let the trial parenthesis be (2). In (1) and (2) we would have a contradiction.

Example 2. Given $p \cdot q \cdot r$; to prove that we have p.

$$\overline{\overline{p\ q\ r}} \qquad \overline{p\ q\ r}$$

(1) (2)

We have (1) by hypothesis. Let the trial parenthesis be (2). In (1) and (2) we would have a contradiction.

Example 3. Given p and $p \supset q$; to prove that we have q.

$$\begin{array}{ccc} p & \overline{\bar{p}\ \bar{q}} & \bar{q} \\ (1) & (2) & (3) \end{array}$$

We have (1) and (2) by hypothesis. Let the trial parenthesis be (3). In (1), (2) and (3) we would have a contradiction.

Example 4. Given p; to prove that we have $\sim(\sim p)$.

$$\begin{array}{ccc} p & \bar{\bar{p}} & \bar{p} \\ (1) & (2) & (3) \end{array}$$

We have (1) by hypothesis. Let the trial parenthesis be (2). But if we had (2) we would have (3), by Example 1. In (1) and (3) we would have a contradiction.

In the proof of Example 4, the transition from (2) to (3) constitutes a *step*. A step, in this calculus, can be thought of as a subsidiary calculation, using as hypothesis (if an hypothesis is needed) one or more propositions already established in the principal calculation. In subsequent proofs, steps will be indicated by the formula "$\cdots \rightarrow \cdots$," where the two groups of dots represent, respectively, the hypothesis and conclusions of the subsidiary calculation.

Example 5. Given p and q; to prove that we have $p \cdot q$.

$$\begin{array}{cccc} p & q & \overline{\overline{p\ q}} & \overline{p\ q} \\ (1) & (2) & (3) & (4) \end{array}$$

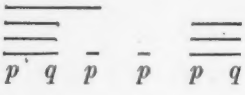
We have (1) and (2) by hypothesis. Let the trial parenthesis be (3). (3) $\rightarrow$ (4). In (1), (2) and (4) we would have a contradiction.

Example 6. Given p; to prove that we have $p \vee q$.

$$\begin{array}{ccc} p & \overline{\bar{p}\ \bar{q}} & \bar{p} \\ (1) & (2) & (3) \end{array}$$

We have (1) by hypothesis. Let the trial parenthesis be (2). (2) $\rightarrow$ (3). In (1) and (3) we would have a contradiction.

Example 7. Given $(p \cdot q) \vee p$; to prove that we have p.

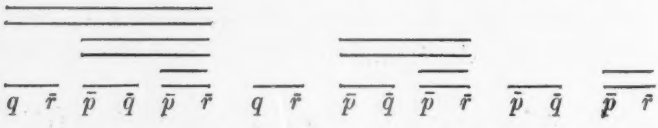


(1) (2) (3)

We have (1) by hypothesis. Let the trial parenthesis be (2). (2) → (3). In (1), (2) and (3) we would have a contradiction.

In Example 7, the step (2) → (3) is a subsidiary calculation which in turn contains subsidiary calculations.

Example 8. To prove that we have $(q \supset r) \supset [(p \vee q) \supset (p \vee r)]$.



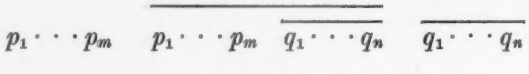
(1) (2) (3) (4) (5)

(6) (7) (8)

Let the trial parenthesis be (1). (1) → (2)(3). (3) → (4)(5). (5) → (6)(7). (2)(7) → (8). In (4), (6) and (8) we would have a contradiction.

It will now be demonstrated that if we have $p_1 \cdot \cdot \cdot p_m$ $\overline{q_1 \cdot \cdot \cdot q_n}$ and $q_1 \cdot \cdot \cdot q_n$ $\overline{p_1 \cdot \cdot \cdot p_m}$ $(m = 1, 2, \cdot \cdot \cdot ;\ n = 1, 2, \cdot \cdot \cdot)$, then $q_1, \cdot \cdot \cdot, q_n$ collectively can be substituted for $p_1, \cdot \cdot \cdot, p_m$ collectively in any formula where $p_1, \cdot \cdot \cdot, p_m$ are not covered or are all covered by the same lines.

To take first the simpler case, assume that $p_1, \cdot \cdot \cdot, p_m$ are not covered. We have $q_1, \cdot \cdot \cdot, q_n$ by calculation as follows:



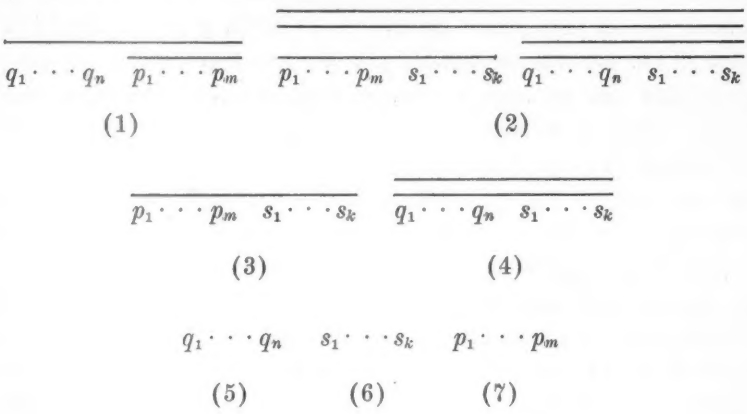
(1) (2) (3)

We have (1) and (2) by hypothesis. Let the trial parenthesis be (3). In (1),

(2) and (3) we would have a contradiction. Hence $q_1, \cdots, q_n$ can be written without being covered, and can be substituted for $p_1, \cdots, p_m$.

But if $p_1, \cdots, p_m$ are covered by j lines ($j = 1, 2, \cdots$) then we find that $p_1, \cdots, p_m$ occur in P_1, P_1 occurs in $P_2, \cdots, P_{j-1}$ occurs in P_j; where $P_1, \cdots, P_j$ are parentheses, and P_j is not covered. Now consider Q_j, obtained from P_j by substituting $q_1, \cdots, q_n$ for $p_1, \cdots, p_m$. We find that $q_1, \cdots, q_n$ occur in Q_1, Q_1 occurs in $Q_2, \cdots, Q_{j-1}$ occurs in Q_j; where $Q_1, \cdots, Q_j$ are parentheses. We also find that P_1 can be represented by $\overline{p_1 \cdots p_m \quad s_1 \cdots s_k}$, and Q_1 by $\overline{q_1 \cdots q_n \; s_1 \cdots s_k}$ ($k = 0, 1, 2, \cdots$), where $s_1, \cdots, s_k$ are the propositions other than $p_1, \cdots, p_m$ which occur in P_1. (In our notation here we take advantage of the circumstance that propositions within a parenthesis can be written in any order without affecting the calculus.)

We have $\overline{P_1 \quad \overline{Q_1}}$ by calculation as follows:



We have (1) by hypothesis. Let the trial parenthesis be (2). (2) → (3)(4). (4) → (5)(6). (1)(5) → (7). In (3), (6) and (7) we would have a contradiction. Similarly we have $\overline{Q_1 \quad \overline{P_1}}$.

Again, let i be any integer $2, 3, \cdots, j$. If we have $\overline{P_{i-1} \quad \overline{Q_{i-1}}}$ and $\overline{Q_{i-1} \quad \overline{P_{i-1}}}$, then we have $\overline{P_i \quad \overline{Q_i}}$ and $\overline{Q_i \quad \overline{P_i}}$ by a calculation similar to that in the foregoing paragraph. (Let P_{i-1} take the place of $p_1, \cdots, p_m$; Q_{i-1} the place of $q_1, \cdots, q_n$; P_i the place of P_1; Q_i the place of Q_1.) But we have $\overline{P_1 \quad \overline{Q_1}}$ and $\overline{Q_1 \quad \overline{P_1}}$; therefore by induction we have $\overline{P_j \quad \overline{Q_j}}$ and $\overline{Q_j \quad \overline{P_j}}$.

Since we have P_j and $P_j \ \overline{Q_j}$, we have Q_j by Example 3. Hence Q_j can be written without being covered; in effect, $q_1, \cdots, q_n$ can be substituted for $p_1, \cdots, p_m$ in P_j.

Thus in all cases, $q_1, \cdots, q_n$ can be substituted for $p_1, \cdots, p_m$. Similarly $p_1, \cdots, p_m$ can be substituted for $q_1, \cdots, q_n$.

The principle demonstrated above can be used to shorten a great many calculations. We have $\overline{\overline{p_1 \cdots p_m} \ \overline{p_1 \cdots p_m}}$ and $p_1 \cdots p_m \ \overline{\overline{p_1 \cdots p_m}}$ by calculation; therefore $p_1, \cdots, p_m$ collectively can be substituted for $\overline{p_1 \cdots p_m}$; in effect, *two lines of the same length, one above the other, cancel each other.* Thus in Example 8 we might go from (1) to (2), (4), (6) and (7) simply by cancelling pairs of lines in (1).

Another rule for cancellation can be stated as follows: *If we have a proposition, p, and have also a parenthesis in which p occurs, then p may be cancelled in the parenthesis (unless p is the* ONLY *proposition in the parenthesis—in that case we have a contradiction, and the calculation is already complete).* Thus in Example 8, by referring to (7) we might cancel the $\bar{r}$ in (2), making (2) equivalent to (8).

We see that the entire proof of Example 8 could be accomplished by cancellations in the trial parenthesis. The same may be said of any theorem in the first three chapters of *Principia Mathematica.*

It remains to be demonstrated that the Calculus of Non-contradiction is not weaker than the calculus of propositions in *Principia Mathematica.* Examples 9, 10, 11 and 12, below, and Example 8, above, correspond to the five axioms of *Principia*—propositions *1.2, *1.3, *1.4, *1.5 and *1.6 respectively. In Example 3 we have the principle of inference. In Examples 13 to 20 we have $D \supset D'$ and $D' \supset D$, where D is the definiendum and D' the definiens, for each of the following definitions in *Principia*: *1.01, *2.33, *3.01, *4.34. It is evident, then, that we can prove any theorem which can be proved by the method of *Principia Mathematica,* for if necessary we can imitate that method.

The proofs of Examples 9 to 20 can be accomplished, in each instance, by cancellations in the trial parenthesis, using only the first rule for cancellation.

Example 9. To prove that we have $(p \vee p) \supset p$.

Let the trial parenthesis be $\overline{\bar{p} \ \bar{p} \ \bar{p}}$.

Example 10. To prove that we have $q \supset (p \vee q)$.

Let the trial parenthesis be $q \quad \bar{p} \quad \bar{q}$.

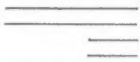
Example 11. To prove that we have $(p \vee q) \supset (q \vee p)$.

Let the trial parenthesis be $\bar{p} \quad \bar{q} \quad \bar{q} \quad \bar{p}$.

Example 12. To prove that we have $[p \vee (q \vee r)] \supset [q \vee (p \vee r)]$.

Let the trial parenthesis be $\bar{p} \quad \bar{q} \quad \bar{r} \quad \bar{q} \quad \bar{p} \quad \bar{r}$.

Example 13. To prove that we have $(p \supset q) \supset (\sim p \vee q)$.

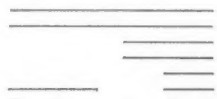


Let the trial parenthesis be $p \quad \bar{q} \quad \bar{p} \quad \bar{q}$.

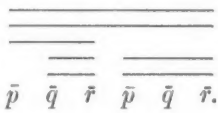
Example 14. To prove that we have $(\sim p \vee q) \supset (p \supset q)$.

Let the trial parenthesis be $\bar{p} \quad \bar{q} \quad p \quad \bar{q}$.

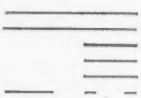
Example 15. To prove that we have $(p \vee q \vee r) \supset [p \vee (q \vee r)]$.



Let the trial parenthesis be $\bar{p} \quad \bar{q} \quad \bar{r} \quad \bar{p} \quad \bar{q} \quad \bar{r}$.

Example 16. To prove that we have $[p \vee (q \vee r)] \supset (p \vee q \vee r)$.

Let the trial parenthesis be $\bar{p} \quad \bar{q} \quad \bar{r} \quad \bar{p} \quad \bar{q} \quad \bar{r}$.

Example 17. To prove that we have $(p \cdot q) \supset [\sim (\sim p \vee \sim q)]$.

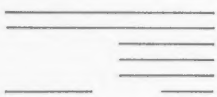


Let the trial parenthesis be $p \quad q \quad p \quad q$.

Example 18. To prove that we have $[\sim (\sim p \vee \sim q)] \supset (p \cdot q)$.

Let the trial parenthesis be $p \quad q \quad p \quad q$.

Example 19. To prove that we have $(p \cdot q \cdot r) \supset [p \cdot (q \cdot r)]$.



Let the trial parenthesis be $p \quad q \quad r \quad p \quad q \quad r$.

Example 20. To prove that we have $[p \cdot (q \cdot r)] \supset (p \cdot q \cdot r)$.

Let the trial parenthesis be $p \quad q \quad r \quad p \quad q \quad r$.

208 Howe Street,
Seattle, Washington.

SPLITTINGS OF A LATTICE.*

By PHILIP M. WHITMAN.

In another paper ([11] Lemma 1.1) the author has shown that in $FL(n)$—*i. e.*, the free lattice generated by $x_1, \cdots, x_n$; see also [10]—when n is finite, given any lattice polynomial b and index p, then either $x_p \leq b$ or $b \leq V_{i \neq p} x_i$, the two cases being mutually exclusive. This is a special case of the general situation considered in the present paper: we say a lattice is split if it is partitioned into two classes, an ideal and a dual ideal (detailed definitions are given below). Birkhoff and Stone have investigated this situation to some extent (cf. **3**), chiefly showing that any distributive lattice can be split in certain ways. Ward and Dilworth ([9] p. 332) have generalized the notion in the case of residuated lattices. In the present paper, we consider the properties of splittings and successions of splittings in general lattices.

In **2** we study their existence in terms of relations between the generating elements of the lattice. The relation of splittability to modularity and other properties is discussed in **3** and part of **5**. In **4** we make definite the intuitive notion of the "edge" along which the splitting seems geometrically to take place; under suitable conditions the two partition classes ("splinters") have isomorphic edges. In **5** we reverse the process of splitting and study the conditions under which two lattices may be combined to form a single lattice of which they are splinters. Here too we find that a principally split lattice is modular (distributive) if and only if each splinter has this property and the edges are convex in the splinters. In **7** we observe that the splinters from a splitting may themselves be splittable, leading to a discussion of successive splittings—or chippings, as they may appropriately be called.

For guidance in the preparation of this paper, the author is indebted to Prof. S. MacLane in the earlier stages and Prof. G. Birkhoff in the later stages.

1. Definitions and notation. For general definitions and notation in lattice theory we refer the reader to Birkhoff's book [3]; for instance $\leq$ or $\geq$ denotes the order relation, $\cap$ and $\cup$ the meet and join of two elements, Λ and V the meet and join of more than two. $A + B$ is the set union of A and B, while $\subset$ means set inclusion. We may recall some pertinent specialized definitions:

* Received September 13, 1941; Presented in part to the American Mathematical Society September 10, 1940.

DEFINITION 1.1. In a lattice, an element a **covers**[1] an element b, denoted $a \underline{>>} b$, if $a > b$ but for no c is $a > c > b$.

DEFINITION 1.2. A **lattice polynomial** in $x_1, \cdots, x_n, \cdots$ is any finite combination of the x_i by $\cap$ and $\cup$, repetitions being allowed, e. g.,

$$(x_1 \cup x_1) \cap \{x_2 \cap [x_3 \cup (x_4 \cap x_1)]\}.$$

Cf. [3], p. 21.

DEFINITION 1.3. A **lattice** L **is generated** by elements $x_1, \cdots, x_n, \cdots$ if L consists of all lattice polynomials in the x_i.

The set of all elements of the lattice makes up one set of generators, but there may be much smaller sets. For instance (Birkhoff [1] p. 463-464, [2] p. 451; see also the author, [10] p. 330, [11] § 5), there are lattices with only three generators which have infinitely many elements.

DEFINITION 1.4. The **length** of a lattice polynomial b, denoted by $L(b)$, is the number of generators appearing in it, counting repetitions.

Thus $L(x_1) = 1$, $L([x_1 \cup x_1] \cap x_2) = 3$. Note that $L(b) = 1$ if and only if b is a generator. It may of course happen that $a = b$, yet $L(a) \neq L(b)$; the length is concerned only with the formal polynomial. This definition (cf. [10] p. 325, 326) is of technical use in proofs by induction; it of course presupposes a choice of generators.

DEFINITION 1.5. An **ideal** J in a lattice L is a set of elements of L such that $a \in J$, $b \in J$, and $c \in L$, imply $a \cup b \in J$ and $a \cap c \in J$. An ideal is **principal** if it is the set of all elements of L which are contained in (i. e., $\leq$) some fixed element p_0; we say that p_0 **generates the ideal**. An ideal J is **prime** if $a \cap b \in J$ implies $a \in J$ or $b \in J$. **Dual** prefixed to any of these means that $\cap$ and $\cup$, $\leq$ and $\geq$ are to be interchanged.[2]

Two equivalent properties ([3] p. 78) will illuminate the situation:

LEMMA 1.1. *An ideal J in a lattice L is prime if and only if the set complement of J in L (denoted by $L - J$) is a dual ideal.*

LEMMA 1.2. *An ideal J in a lattice L is prime if and only if J is the set of antecedents of 0 under some homomorphism of L on the two element lattice $[0 < I]$.*

DEFINITION 1.6. A lattice L is **split** if it is partitioned into an ideal J' and a dual ideal J'', each non-void; it is **principally split** if J' and J'' are principal.

---

[1] Some authors say, "is prime over."

[2] Dilworth [5] interchanges "ideal" and "dual ideal."

The appropriateness of the term "split" (Birkhoff [4] speaks of a "cleavage"; Ward and Dilworth, [9] p. 332, have used "cleavage" in extending the idea in a somewhat different direction) lies in a rough analogy with the case of a log or timber placed on its side and diagonally split. We call J' and J'' the bottom and top **splinters** of L, and the partition a **splitting** of L. If J' and J'' are principal we frequently denote them by $\boldsymbol{P}$ and $\boldsymbol{D}$ and the elements generating them by $\boldsymbol{p}_0$ and $\boldsymbol{d}_0$, and say that p_0 and d_0 split L. There exist lattices which cannot be split; e. g., the characteristic five-element modular non-distributive lattice (Birkhoff [3] p. 75, Corollary 2). On the other hand, the lemma cited in the introduction shows that $FL(n)$ can be split, for given p; then the set J' of all b such that $b \leq V_{i \neq p} x_i$ constitutes a principal ideal and the set J'' of those such that $x_p \leq b$ forms a dual principal ideal. By the lemma cited, J' and J'' are disjoint and exhaust $FL(n)$ as required by Definition 1.6; we may also conclude by Lemma 1.1 that each is prime.

2. Existence theorems. In view of the preceding remarks, we seek conditions under which splittings will exist. For clarity we first state the results in case the lattice has a finite number of generators (Theorem 1) and then state and prove (Theorem 1′) a more general theorem. We observe that if a lattice L is split, the generators are thereby partitioned into two classes $\{x_j\} \subset J'$ and $\{y_i\} \subset J''$, and for all finite sets R and T, $\Lambda_{i \in R} y_i \in J''$, $V_{j \in T} x_j \in J'$, so that

$$\Lambda_{i \in R} y_i \not\leq V_{j \in T} x_j.$$

It turns out that we may essentially reverse this statement.

THEOREM 1. *Let L be a lattice with a finite number of generators, and let the generators be exhausted by two (possibly overlapping) non-void sets $\{x_1, \cdots, x_r\}$ and $\{y_1, \cdots, y_s\}$. Then, given any element b of L, either* (1) $b \leq \overset{i=r}{\underset{i=1}{V}} x_i$ *or* (2) $\overset{j=s}{\underset{j=1}{\Lambda}} y_j \leq b$. *If moreover* $\overset{j=s}{\underset{j=1}{\Lambda}} y_j \not\leq \overset{i=r}{\underset{i=1}{V}} x_i$ *(then a fortiori $\{x_i\}$ and $\{y_j\}$ are disjoint) then the partition of L into the sets J' and J'' satisfying* (1) *and* (2) *respectively is a principal splitting of L.*

THEOREM 1′. *Let the generators of a lattice L be exhausted by two non-void sets $\{x_i\}$ and $\{y_j\}$. Given any element b of L, then* (A) *either* (1) $b \leq V_{i \in R} x_i$ *for some finite set R, or else* (2) $\Lambda_{j \in T} y_j \leq b$ *for some finite set T. Moreover* (B) *the set J' of all b for which* (1) *holds is an ideal and the set J'' of all b for which* (2) *holds is a dual ideal.* (C) *If for all finite sets R and T, $\Lambda_{j \in T} y_j \not\leq V_{i \in R} x_i$, then J' and J'' are disjoint and L is split into them.*

(D) *If besides the conditions of* (C), *also* $p_0 = V_{all\ i} x_i$ *and* $d_0 = \Lambda_{all\ j} y_j$ *exist and are in* J' *and* J'' *respectively* (e. g., *if the number of generators is finite*) *then* J' *and* J'' *are principal.*

Proof. That either (1) or (2) must hold is obvious for $L(b) = 1$; we proceed by induction on $L(b)$. If $b \equiv Vb_k$ and the statement were not true of b, then for all finite sets T, $\Lambda_{j \epsilon T} y_j \not\leq b$. Hence, given k, then $\Lambda_{j \epsilon T} y_j \not\leq b_k$ for all finite T. By induction, given k, $b_k \leq V_{i \epsilon R_k} x_i$ for some finite R_k. Consequently $b \leq V_{i \epsilon U} x_i$, where U is the set union of the R_k and so is again a finite set as desired. The case $b \equiv \Lambda b_k$ is dual. Thus (A) is proved. The other statements are then obvious.

Since in the free lattice $FL(n)$ the conditions of Theorem 1′C hold by [10] Corollary (17) or [11] Theorem 1, we have:

COROLLARY 2.1. *Any partition of the generators of* $FL(n)$ *into two non-void classes determines a splitting of* $FL(n)$ *with the elements of the first class in the bottom splinter. Hence* $FL(n)$ *has* $2^n - 2$ *splittings.*

Theorem 1 specializes Theorem 1′ to the case of a finite number of generators, yielding principal splittings. But we can discuss principal splittings even if there are infinitely many generators:

THEOREM 2. $p_0 = V_{all\ i} x_i$ *and* $d_0 = \Lambda_{all\ j} y_j$ *principally split a lattice* L *generated by the* x_i *and* y_j *if and only if they exist in* L *and* $d_0 \not\leq p_0$. *Whether or not these conditions hold, there are no other principal splittings of* L. *If the conditions do hold, then* $p_0 = V_{i \epsilon R} x_i$ *and* $d_0 = \Lambda_{j \epsilon T} y_j$ *for some finite sets* R *and* T.

Proof. In the first sentence, "if" follows from Theorem 1′, while "only if" is required by Definition 1. 6. Now consider any principal splitting, say by p and d into $P \equiv J'$ and $D \equiv J''$. Let $\{x_i\}$ be the generators of L which lie in P. By hypothesis $x_i \leq p$ for all i. But by the remarks at the beginning of **2**, the first condition in Theorem 1′ is satisfied, so $p \leq V_{k \epsilon R} x_k$ for some finite R. Thus $x_i \leq V_{k \epsilon R} x_k$ for all i, so $V_{k \epsilon R} x_k$ is the least upper bound of the x_i, and

$$p_0 = V_{all\ i} x_i = V_{i \epsilon R} x_i = p$$

and dually in D. Thus the splitting is of the form specified, and so are p_0 and d_0.

d_0 and p_0 will exist, for instance, if the number of generators is finite or if the finite chain conditions hold or if J' and J'' are complete ([3] p. 17), but in general they need not, since L by definition consists of the *finite* combinations of the generators. For instance, they fail to exist in a chain of

ordinal type $\omega + \omega^*$ partitioned into splinters of type ω and ω^*. We may extend § 4 of [10] and Theorem 3 of [11] slightly:

COROLLARY 2.2. *In* $FL(n)$ *for* $k < n < \infty$,

$$(\overset{k}{\underset{i=1}{V}} x_i) \cup (\overset{n}{\underset{i=k+1}{\Lambda}} x_i) \quad \text{covers} \quad \overset{k}{\underset{i=1}{V}} x_i.$$

3. Relation of splittability to other properties. Splittability and modularity are completely independent.[3] Sometimes both hold, as with free modular lattices with more than one generator, especially that with two generators (4 elements), sometimes only splittability, as with free lattices with more than two generators, sometimes only modularity, as with the typical five-element modular non-distributive lattice ([3] p. 75), and sometimes neither as is the case if an extra element is interpolated in one of the chains of the lattice just cited. Likewise suitable examples show that splittability and the existence of complements are completely independent.

On the other hand any distributive lattice with more than one element can be split. This was shown by Birkhoff, [1] Theorem 21.1 and [4] p. 449, and Stone, [8] Theorems 6-7. In fact, they show that given elements a and b such that $a \not\leq b$, then there is a splitting with $a \in J''$ and $b \in J'$. A much simpler proof[4] can however be given if we wish to show only the splittability.

4. Edges of the splinters. The geometrical interpretation of splitting pointed out in **1** suggests that the imaginary "line" along which the splitting takes place should in some sense represent an edge of the two splinters. For principal splittings, a natural sense seems to be this: **in this section, we suppose a lattice L is principally split into P and D by p_0 and d_0,** and make the following

DEFINITION 4.1. An element a of D is in the **edge** of D, denoted $a \in E(D)$, if a covers some element of P; dually for the edge of P. The edge of the splitting is the set union of $E(D)$ and $E(P)$.

The following lemma assures us that the edges as thus defined really exist—that there is not a pair of infinite chains of elements approaching the intuitive line of cleavage, one from each side, so as to avoid having any element of one cover any element of the other.

---

[3] In the sense that the presence or absence of one neither assures nor forbids that of the other. See however Theorems 5 and 6.

[4] Not published. See the author's thesis, Harvard University, 1940, p. 58.

LEMMA 4.1. *If $a \in D$ and $b \in P$, and $b \leq a$, then*

$$b \leq (b \cup d_0) \cap p_0 \underline{\ll} b \cup d_0 = [(b \cup d_0) \cap p_0] \cup d_0 \leq a.$$

Thus $b \cup d_0 \in E(D)$ and $(b \cup d_0) \cap p_0 \in E(P)$. Also dually.

Proof. $b \in P$ implies $b \leq p_0$ by Definition 1.6, so $b \leq (b \cup d_0) \cap p_0$. Likewise $b \cup d_0 \leq a$. Obviously

$$[(b \cup d_0) \cap p_0] \cup d_0 \leq b \cup d_0 \leq [(b \cup d_0) \cap p_0] \cup d_0.$$

As for the covering relation stated, we have

$$(b \cup d_0) \cap p_0 < b \cup d_0$$

(equality cannot hold since they are in different splinters), and if

$$(1) \qquad (b \cup d_0) \cap p_0 < c < b \cup d_0$$

then either $c \in P$ or $c \in D$. In the former case however, $c \leq p_0$ so $c \leq (b \cup d_0) \cap p_0$, a contradiction, and in the latter case $d_0 \leq c$ and $b \leq p_0$ so by (1),

$$b \cup d_0 = [(b \cup d_0) \cap p_0] \cup d_0 \leq c \cup d_0 = c,$$

contrary to the sharp inequality in (1). Thus the covering relation is proved, and the rest of the lemma follows immediately.

COROLLARY 4.1. *$b \in E(P)$ is covered by exactly one element of $E(D)$, namely by $b \cup d_0$, and dually for $b \in E(D)$.*

LEMMA 4.2. *(i) If a_1 and a_2 are in $E(D)$ so is $a_1 \cup a_2$. The dual refers to $E(P)$. (ii) If $a \in E(D)$ and $b \in P$, then $a \cup b \in E(D)$.*

Proof. (*i*) Say a_i covers $b_i \in E(P)$. Now $b_1 \cup b_2 \in P$ since P is an ideal, so if $a_1 \cup a_2$ is not in $E(D)$ then

$$b_1 \cup b_2 < e < a_1 \cup a_2$$

for some $e \in E(D)$, by Lemma 4.1. Thus, by Corollary 4.1,

$$e < a_1 \cup a_2 = (d_0 \cup b_1) \cup (d_0 \cup b_2) = d_0 \cup (b_1 \cup b_2) \leq e$$

which is impossible. (*ii*) $b \cup d_0 \in E(D)$ by Lemma 4.1, so

$$a \cup b = (a \cup d_0) \cup b = a \cup (d_0 \cup b)$$

is in $E(D)$ by (*i*).

THEOREM 3. *$E(D)$ and $E(P)$ are lattices, though not generally sublattices of D and P, and as such are isomorphic.*

Proof. By Corollary 4.1, their elements are in a natural one-to-one correspondence. $E(D)$ and $E(P)$ are partially ordered by the inclusion relation of L. If a_1 and a_2 in $E(D)$ cover respectively b_1 and b_2 in $E(P)$, and $a_1 \leq a_2$ but $b_1 \not\leq b_2$, then $b_2 < b_1 \cup b_2 \leq a_2$; but $b_1 \cup b_2 \in P$, so $b_1 \cup b_2 < a_2$ and a_2 could not cover b_2. Hence $E(D)$ and $E(P)$ are isomorphic as partially ordered sets. By Lemma 4.2, any two elements of $E(D)$ have a least upper bound in $E(D)$, namely the same as in L. Dually, any pair of elements of $E(P)$ have a greatest lower bound, so by the isomorphism as partially ordered sets, greatest lower bounds exist in $E(D)$, which is therefore a lattice. Dually for $E(P)$.

NOTE 4.1. d_0 is the zero of $E(D)$, $d_0 \cup p_0$ the unit, and dually for $E(P)$. That the edges, though lattices, need not be sublattices of D and P (nor of L) is readily shown by an example.

COROLLARY 4.2. *If $a \in D$, $c \in E(D)$, and $c \leq a$, then $c \leq (a \cap p_0) \cup d_0$ which is in $E(D)$.*

Proof. By Lemma 4.1, if c covers $c' \in E(P)$, then $c' \leq a \cap p_0$; by the isomorphism, $c \leq (a \cap p_0) \cup d_0$.

5. Combination of splinters. Having considered the existence of principal splittings and the properties of their edges, we turn now to the converse problem: given two lattices P and D, what about putting them together to get a lattice of which they are splinters?

DEFINITION 5.1. Let there be given disjoint partially ordered sets A and B, and subsets $E(A)$ and $E(B)$ thereof, and a correspondence (not necessarily one-to-one) between the elements of $E(A)$ and $E(B)$. Extend the inclusion relation of A and B to the set union $A + B$ by setting $c_1 \prec c_2$ if and only if

$c_1 \in A$, $c_2 \in A$, and $c_1 \leq c_2$ in A, or
$c_1 \in B$, $c_2 \in B$, and $c_1 \leq c_2$ in B, or

$c_1 \in A$, $c_2 \in B$, and there exist $e_1 \in E(A)$ and $e_2 \in E(B)$ such that $c_1 \leq e_1$ in A, $e_2 \leq c_2$ in B, and e_1 corresponds to e_2 under the given correspondence.

LEMMA 5.1. *Under $\prec$, $A + B$ is a partially ordered set, and $\prec$ and $\leq$ coincide between elements of A, likewise between elements of B.*

Proof. Note that $b \prec a$ is impossible if $a \in A$ and $b \in B$. It then follows from the construction that the postulates for partial ordering ([3] p. 5) hold and that $\prec$ coincides with $\leq$ in A (and in B).

THEOREM 4. *Suppose that disjoint lattices D and P and subsets $E(D)$ and $E(P)$ thereof are given, such that*

(1) *$E(D)$ and $E(P)$ are isomorphic partially ordered sets under the inclusion relations of D and P, and*

(2) *D has a zero d_0, and P has a unit p_0, and*

(3) *the zero of D is in $E(D)$, and given $b \in D$ then there is a greatest element c of $E(D)$ such that $c \leq b$ in D, and*

(3′) *the dual of* (3) *holds for $E(P)$.*

Then taking the correspondence which gives the isomorphism as the correspondence of Definition 5.1, the set union $L = P + D$ is a lattice under $\prec$, with top splinter D and bottom splinter P and edges $E(D)$ and $E(P)$.

Conversely, conditions (1)-(3) *are necessary that L be principally split into D and P with those edges.*

Proof. The necessity is shown by Theorem 3, the definition of principal, Note 4.1, and Corollary 4.2. To prove sufficiency, we first observe that L is partially ordered, by Lemma 5.1. To show L a lattice, we need to find a least upper bound for elements a and b.

Case 1. If a and b are both in P: Then $a \cup b$ (their lattice join in P) is an upper bound in L. In fact it is least, for suppose $a \prec e$ and $b \prec e$. If $e \in P$ then $a \cup b \prec e$ by construction and Lemma 5.1. Otherwise $e \in D$. Then of the $f \in E(D)$ for which $f \prec e$ there is by (3) a greatest, say e_1. Since $a \prec e$, then $f_1 \prec e$ for some $f_1 \in E(D)$, where f_1 corresponds under the isomorphism to $f'_1 \in E(P)$ and $a \prec f'_1$. Likewise $b \prec f'_2 \prec f_2 \prec e$ with similar conditions on f_2 and f'_2. By choice of e_1 we have $f_1 \prec e_1$ and $f_2 \prec e_1$. Say e_1 covers $e'_1 \in E(P)$; then by the isomorphism, $f'_1 \prec e'_1$ and $f'_2 \prec e'_1$ and all of these are in P, so $a \cup b \prec e'_1 \prec e$ as desired.

Case 2. a and b both in D. Then any upper bound is in D too, so $a \cup b$ serves also in L.

Case 3. $b \in D$ and $a \in P$. Let a_1 be the least element of $E(P)$ containing a, by (3′), and let it correspond to $a_1'' \in E(D)$. Then $a \prec a_1''$, and if $f \in D$ and $a \prec f$ then $a_1'' \prec f$ by the dual of the construction of e'_1 in Case 1. Then the lattice union in D, $b \cup a_1''$, is the least upper bound of a and b in L, for $b \prec g$ and $a \prec g$ imply $g \in D$ and $a_1'' \prec g$. Thus l. u. b. always exists; by duality L is a lattice.

L can be principally split into D and P, for let d_0 be the zero of D and p_0 the unit of P. By construction $a \in P$ and $b \in D$ imply $b \not\prec a$; hence $d_0 \not\prec p_0$.

Yet given a, then either $a \in D$ or $a \in P$, so $d_0 \prec a$ or $a \prec p_0$. This completes the proof.

Thus the top and bottom splinters and edges completely determine the lattice unless the edges admit automorphisms; in the latter case, if one correspondence of the edges yields a lattice, so does any correspondence resulting from an automorphism. However, the two lattices obtained may be non-isomorphic, as may be shown by suitable examples.

If S is a subset of L, then S is said ([3] p. 9) to be **convex** in L if whenever a and b are in S, c in L, and $a < c < b$, then c is also in S. If S is a convex set with zero a and unit b then we call S an **interval sublattice**, for it consists of all x in the interval $a \leq x \leq b$, and it will be a sublattice of L. We recall ([3] p. 34) that a lattice is modular if and only if it does not contain a certain five-element sublattice, graphed on page 35 of [3], which we shall call ***NM5***. Its elements obey the relations

$$(i)\ f < e < a \qquad (iii)\ e \cup c = e \cup b = a$$

$$(ii)\ f < c < b < a \qquad (iv)\ e \cap c = e \cap b = f.$$

THEOREM 5. *If a lattice L is principally split into P and D, then L is modular (or distributive, respectively) if and only if*

(I) *P and D are modular (distributive) and*

(II) *the edges are interval sublattices of the respective splinters.*[5]

Proof. Condition (I) is obviously necessary. As for (II), we note that d_0 is the zero of $E(D)$ and $d_0 \cup p_0$ the unit by Note 4. 1. Thus to show (II) necessary, we suppose the contrary, that

$$d_0 < c < d_0 \cup p_0 \text{ but } c \text{ not in } E(D)$$

(or dually). Let $a = c \cap p_0$; then $a \in E(P)$ by Lemma 4. 2 (ii) since $p_0 \in E(P)$. Let a be covered by $a'' \in E(D)$. Then we assert that a, p_0, a'', c, and $d_0 \cup p_0$ form $NM5$. As proof, we see that (i) $a < p_0 < d_0 \cup p_0$, since $a \in P$ implies $a \leq p_0$ and $a = p_0$ would give $d_0 \cup p_0 \leq c$ contrary to hypothesis, while $p_0 \neq p_0 \cup d_0$ since one is in P and one in D. (ii) $a < a'' < c < d_0 \cup p_0$ since $a'' \neq c$ (because one is in $E(D)$ and one is not) and yet $a'' = a \cup d_0 \leq c$. (iii) $p_0 \cup c = p_0 \cup a'' = d_0 \cup p_0$, since

$$d_0 \cup p_0 \leq a'' \cup p_0 \leq c \cup p_0 \leq d_0 \cup p_0$$

by construction. (iv) $p_0 \cap c = p_0 \cap a'' = a$, since

$$a \leq a'' \cap p_0 \leq c \cap p_0 = a$$

---

[5] We may replace "interval sublattices of" by "convex in," thus formally strengthening the sufficiency but weakening the other half of the theorem.

by construction. Thus if (II) fails, L contains $NM5$, so (II) is necessary for modularity, and a fortiori for distributivity.

To show sufficiency, suppose that (I) and (II) hold, but L is not modular. Then it has a sublattice $NM5$, consisting of elements a, b, c, e, f with the relations (i)-(iv). If all five were in D or all in P then (I) would be violated. If four lie in one splinter and one in the other, then the definition of the splinters as ideals is violated (e. g., if $f \in P$ and the others are in D, then $b \in D$, $c \in D$, hence by definition of dual ideal $b \cap c = f$ is in D, a contradiction) and similarly the only possibility is that

(1) $\qquad a, b, c$ are in D; e, f are in P

(or dually). In this case, we shall reach a contradiction to hypothesis (I), via steps (2) to (8) below. By Corollary 4. 2, we may

(2) $\qquad$ let e' be the least element of $E(P)$ such that $e \leq e'$.

By Lemma 4. 1, $e' = (e \cup d_0) \cap p_0$ is covered by $e'' = e \cup d_0$; thus $e'' \leq a$. By construction $e'' \in E(D)$ and

$$d_0 \leq c \cap e'' \leq b \cap e'' \leq e''$$

so by (II), $b \cap e''$ and $c \cap e''$ are in $E(D)$.

(3) Set $b'' = b \cap e''$ and $c'' = c \cap e''$, and suppose that they cover b' and c', in $E(P)$, respectively. $c'' \leq b'' \leq e''$ by construction.

(4) $\qquad b' = b'' \cap e' = b \cap e' \quad$ and $\quad c' = c'' \cap e' = c \cap e'$.

Proof.

$$b' = b' \cap e' \leq b'' \cap e' \leq b'' \cap p_0 = b',$$

the first equality coming from the isomorphism of the edges and the fact that $b'' \leq e''$ by the construction of b''; similarly for c'.

(5) $a = e''$ and $c'' = b''$ cannot both hold.

Proof. If they did, then

$$c = c \cap a = c \cap e'' = c'' = b'' = b \cap e'' = b \cap a = b,$$

whereas $c < b$ by hypothesis.

(6) $c = c''$ implies $a = e''$.

Proof. Then

$$a = c \cup e \leq c \cup e'' = c'' \cup e'' = e'' \leq a.$$

(7) c'', b'', and e'' are all distinct.

Proof. $c'' \leq b'' \leq e''$. If $b'' = e''$ then $e \leq b$ whereas by hypothesis $o < a = e \cup b$, so $b'' \neq e''$. If $c'' = b''$ then by (5), $a \neq e''$, so by (6), $c \neq c''$. Then a, b, c, e'', and $c'' = b''$ form an $NM5$ in D, for

$$a = c \cup e \leq c \cup e'' \leq b \cup e'' \leq a$$

so that (*iii*) $b \cup e'' = c \cup e'' = a$ and (*iv*) $b \cap e'' = c' = c \cap e''$. This contradicts (I). Hence $c'' < b'' < e''$ and

(8) $c' < b' < e'$, by the isomorphism of the edges. We proceed to show e', b', c', e, and f an $NM5$ in P.

(*i*) $f < e < e'$ and (*ii*) $f < c' < b' < e'$.

Proof. $f < e \leq e'$ and $f \leq c' < b' < e'$ by (4) and (8). If $e = e'$ then $f = c \cap e = c \cap e' = c'$ by (4). Conversely if $f = c'$ then $e = e'$ by (II), so if $e = e'$ or $f = c'$ then both hold and

$$c' = f = b \cap e = b \cap e' = b'$$

contrary to (8). Thus the inequalities are sharp as stated.

(*iii*) $b' \cup e = c' \cup e = e'$.

Proof. All are in $E(P)$ by (II) and contain e, and e' is the least such by choice, whereas $c' \cup e \leq b' \cup e \leq e'$.

(*iv*) $b' \cap e = c' \cap e = f$, for otherwise $f < b' \cap e \leq b \cap e$ whereas $b \cap e = f$ by hypothesis.

Thus we have $NM5$ in P contrary to (I), and so the theorem is proved in the modular case. It remains to show the sufficiency if P and D are distributive. Then L is at least modular, as we have just shown, and if not distributive it must have the standard five-element modular non-distributive sublattice of [3] p. 75, first part of Figure 10:

$$a \cup b = a \cup c = b \cup c = e, \qquad a \cap b = a \cap c = b \cap c = f.$$

But as in the first part of the present sufficiency proof, such a configuration must by definition of ideal have all its elements in P or all in D, contrary to the hypothesis that these are distributive. The proof is complete.

COROLLARY 5.1. *If the edges admit automorphisms, and if L, as constructed in Theorem* 4, *is modular or distributive for one correspondence, then it is so for all correspondences for which the automorphism holds.*

NOTE 5.1. If the edges of a principal splitting are convex in the splinters, then the edges are sublattices of the splinters. For the convexity and Note 4.1 imply that if a and b are in an edge so are $a \cap b$ and $a \cup b$.

LEMMA 5.2. *If a complemented lattice L is principally split, then the zero and unit of L are in the edges.*

Proof. d_0 has a complement d'_0. Now $d'_0 \in P$, for if $d'_0 \in D$ then $d_0 \cap d'_0 = 0 \in D$ by Definition 1.6, and P would be void. Hence by Lemma 4.2 (*ii*), $d_0 \cup d'_0 = I \in E(D)$. Dually $0 \in E(P)$.

THEOREM 6. *If a complemented modular lattice L can be principally split then L can be partitioned into two isomorphic interval sublattices.*

Proof. Theorem 5 (II) and Lemma 5.2 imply $D \equiv E(D)$ and $P \equiv E(P)$. Then Theorem 3 completes the proof.

Remark. We observe that by **3** every finite Boolean algebra satisfies the conditions of Theorem 6. In this case however the theorem is easily obtainable from the representation by point sets ([3] p. 89). In connection with complementation, we may add

THEOREM 7. *In the free lattice $FL(n)$, for $2 < n < \infty$, no element other than 0 and I has a complement.*

Proof. Suppose the contrary:

$$(1)\ a \cup b = I = \overset{n}{\underset{i=1}{V}} x_i; \qquad (2)\ a \cap b = 0 = \overset{n}{\underset{i=1}{\Lambda}} x_i.$$

From [10] Corollary (17) and from hypothesis (1) it follows that, given i, either $x_i \leq a$ or $x_i \leq b$; likewise, given i, either $a \leq x_i$ or $b \leq x_i$. Say $x_1 \leq a$. *Case* 1: $a \leq x_1$. Then $x_1 = a$. But $x_i \leq x_j$ if and only if $i = j$, so for $i \neq 1$ we must have $x_2 \leq b$ and $b \leq x_2$. Thus $x_2 = b = x_3$, a contradiction. *Case* 2: $b \leq x_1$. Then $b \leq x_1 \leq a$, so $I = a \cup b = a$ and $0 = a \cap b = b$.

6. Non-principal splittings. We seek now to generalize the results of the preceding sections to the case of non-principal splittings, though in doing so we must sacrifice some of the geometrical interpretation which inspired the nomenclature. Suppose the lattice L is split into J' and J''. Let $\mathfrak{J}'$ stand for the set of all ideals[6] in the bottom splinter J', $\mathfrak{J}''$ for the set of all dual ideals in the top splinter J''. We observe that in particular, if S is any fixed subset of the elements of J'', then the set of all elements of J' which are contained in every element of S is an ideal and hence is an element of $\mathfrak{J}'$.

[6] We may recall in this connection Dilworth's result ([5] Theorem 2.1): In a lattice L, let B be an ideal and (a) a principal ideal such that $B \leq (a)$ and $B \neq (a)$; then there exists an ideal C such that $B \leq C$ and (a) covers C in the lattice of all ideals of L. At first sight this would seem what was needed to generalize the idea of edges, but the author has had no success in this direction.

DEFINITION 6.1. F is in the **pseudo-edge** of J', denoted $F \epsilon \mathfrak{E}(J')$, if F is the set of all elements of J' which are contained in every element of some (not necessarily proper) subset S of the elements of J''. Then we write $F = \mathfrak{J}'(S)$. $\mathfrak{E}(J'')$ is defined dually.

In view of the remarks above, we have these relations between sets: $\mathfrak{E}(J') \subset \mathfrak{J}'$ and $\mathfrak{E}(J'') \subset \mathfrak{J}''$. It is readily observed that

COROLLARY 6.1. $\mathfrak{J}(\mathfrak{J}''(T)) = T$ *if and only if* $T \epsilon \mathfrak{E}(J')$; *that is, if* $T = \mathfrak{J}(S)$ *for some* S. *Also dually.*

We consider now the special closure operations $X \rightarrow (X^*)\dagger$ and $Y \rightarrow (Y\dagger)^*$ discussed by Birkhoff ([3] p. 24-25). He begins with a dyadic relation ρ defined between the members of any two abstract aggregates "I"[7] and "I^*". If X is any subset of "I", then he denotes by X^* the set of $x^* \epsilon$ "I^*" such that $x \rho x^*$ for all $x \epsilon X$; reciprocally, if Y is a subset of "I^*", $Y\dagger$ denotes the set of $x \epsilon$ "I" such that $x \rho y$ for all $y \epsilon Y$. In particular we may take J' as "I" and J'' as "I^*", and the inclusion relation of L as ρ. Then in the notation of the present section, X^* is $\mathfrak{J}''(X)$ and $Y\dagger$ is $\mathfrak{J}'(Y)$. By Birkhoff's Theorem 2.11, we may conclude that the lattices of subsets closed under the closure operations named above are dually isomorphic. What are these closed sets in the present case? By Corollary 6.1, they are precisely the members of $\mathfrak{E}(J')$ and $\mathfrak{E}(J'')$. This is reminiscent of Theorem 3. We may do even better: in the lattices of closed sets in Birkhoff's theorem, the inclusion relation is set inclusion; suppose we reverse the order relation in $\mathfrak{J}''$.

DEFINITION 6.2.[8] If F and G are in $\mathfrak{J}'$ and $F \subset G$, then $F \prec G$. If F and G are in $\mathfrak{J}''$ and $G \subset F$, then $F \prec G$.

Thus the dual isomorphism becomes an isomorphism, so

THEOREM 3'. $\mathfrak{E}(J')$ *and* $\mathfrak{E}(J'')$ *are isomorphic lattices under* $\prec$.

In case J' and J'' are principal (and then in accordance with earlier practice would ordinarily be denoted by P and D) we may apply this to prove Theorem 3, as follows:

LEMMA 6.1. *If* J' *and* J'' *are principal then each* F *which is in* $\mathfrak{E}(J')$ *or* $\mathfrak{E}(J'')$ *is principal, and its generating element is in* $E(J')$ *or* $E(J'')$ *respectively. Conversely, if* J' *and* J'' *are principal and* $f' \epsilon E(J')$ *is covered by* $f'' \epsilon E(J'')$, *then* $\mathfrak{J}(\{f''\})$ *is generated by* f'.

---

[7] Not to be confused with the unit element of a lattice. Also "I^*" is not "I"* in the sense of his next sentence.

[8] The reader will observe below that the order relation of Definition 5.1 and Theorem 4 as applied to ideals is an extension of the relation in Definition 6.2.

Proof. Then d_0 and p_0 (as used in **2-5**) exist. Suppose $F = \mathfrak{F}'(S)$, s is any element of S, and J' is generated by p_0. Then F is the principal ideal generated by $p_0 \cap s$, for $p_0 \cap s \in F$ by definition, and if $f \in F$ then $f \leq s$ and $f \leq p_0$ so $f \leq p_0 \cap s$. Also $p_0 \cap s$ is in $E(J')$ by the dual of Lemma 4.2 (*ii*). The converse part of the lemma is obvious.

NOTE 6.1. If J' and J'' are principal then $F \prec G$ for elements of $\mathfrak{E}(J')$ generated by f and g if and only if $f \leq g$; likewise for $\mathfrak{E}(J'')$. Thus we may obtain Theorem 3 as a corollary of Theorem 3′ and Lemma 6.1, though if the details of the proof of Theorem 3′ ([3] p. 24-25) are taken into account the proof is no shorter.

NOTE 6.2. The ideals of any lattice form a complete lattice[9] under set inclusion, for their set intersection will serve as meet, and the ideal generated by their set union as join. This seems well known though seldom stated; Dilworth ([5] p. 329) essentially gives it (the set of all ideals in a lattice does have a unit, namely the whole lattice). We observe however that (unlike the ideals of a ring [3] p. 36) the ideals of a lattice need not form a modular lattice; this is shown by the ideals of the lattice $NM5$ used in **5**.

We now consider the generalization of Theorem 4, starting not with a splitting of a lattice but rather with two lattices which we seek to exhibit as splinters. We have at once, using the correspondence between the pseudo-edges given in Corollary 6.1,

LEMMA 6.2 (5.1′). *If the construction of Definition* 5.1 *is applied to* $\mathfrak{F}'$, $\mathfrak{F}''$, $\mathfrak{E}(J')$, *and* $\mathfrak{E}(J'')$ *then the inclusion relations of* $\mathfrak{F}'$ *and* $\mathfrak{F}''$ *are extended to the set union* $\mathfrak{L} = \mathfrak{F}' + \mathfrak{F}''$ *which is thereby partially ordered. If in* $J' + J''$ *we set* $a \prec b$ *when the elements of* $\mathfrak{L}$ *generated by* a *and* b *are in this relation, then* $L = J' + J''$ *is partially ordered and is isomorphic to* $\mathfrak{L}$ *or a subset thereof.*

THEOREM 4′. *If conditions* (1) *and* (3)-(3′) *of Theorem* 4 *hold for* $\mathfrak{F}''$ *and* $\mathfrak{F}'$ (*the sets respectively of dual ideals in a lattice* J'' *and of ideals in a lattice* J') *and subsets* $\mathfrak{E}(J'')$ *and* $\mathfrak{E}(J')$ *thereof, then the set union* $\mathfrak{L} = \mathfrak{F}' + \mathfrak{F}''$ *is a lattice under the construction of Definition* 5.1, *taking* $\leq$ *of that definition as the* $\prec$ *of Definition* 6.2.

Proof. $\mathfrak{F}'$ and $\mathfrak{F}''$ are principal with J' and J'' in the rôles of p_0 and d_0, and they are lattices by Note 6.2. Then Theorem 4 applied to $\mathfrak{F}'$ and $\mathfrak{F}''$ gives the result stated.

---

[9] In the sense of having unrestricted meets and joins.

COROLLARY 6.2. *If the conditions of Theorem 4′ are fulfilled and each element of* $\mathfrak{E}(J')$ *and* $\mathfrak{E}(J'')$ *is principal then* L *as constructed in Lemma* 6.2 *is a lattice—the same lattice as obtained by applying Theorem* 4 *to* $J' + J''$ *directly, with suitable choice of* $E(J')$ *and* $E(J'')$.

Proof. Then in particular J' and J'' are principal. Take for $E(J')$ and $E(J'')$ the generators of the elements of $\mathfrak{E}(J')$ and $\mathfrak{E}(J'')$. Then Theorem 4 is applicable to $J' + J''$, and $a \prec b$ from Theorem 4 if and only if $a \prec b$ from Lemma 6.2.

THEOREM 4″. *If a lattice* L *and a splitting thereof into* J' *and* J'' *are given, then the conditions of Theorem* 4′ *are satisfied.* L *is isomorphic to the sublattice of* $\mathfrak{L}$ *consisting of the elements of* $\mathfrak{L}$ *generated by the elements of* L.

Proof. That condition (1) of Theorem 4′ holds follows from Theorem 3′. As for (3), suppose that there is given $B \in \mathfrak{J}''$. Then $\mathfrak{J}''(\mathfrak{J}'(B))$ serves as the desired C. Also $J'' \in \mathfrak{E}(J'')$, for $J'' = \mathfrak{J}''(\mathfrak{J}'(J''))$.

It remains to show the stated isomorphism; that is, if a and b in L generate A and B in the lattice $\mathfrak{L}$ of Theorem 4′, then $a \cap b$ generates $A \cap B$ and dually. *Case* 1: A and B both in $\mathfrak{J}'$. Then by Note 6.2, $A \cap B$ is just their set intersection. But $a \cap b$ is in $A \cap B$ since A and B are ideals, and if $c \in A \cap B$ then $c \in A$ and $c \in B$ so $c \leq a$, $c \leq b$, and $c \leq a \cap b$. Thus $A \cap B$ is generated by $a \cap b$. *Case* 2: A and B both in $\mathfrak{J}''$. Then by Theorem 4′ their meet $A \cap B$ in $\mathfrak{J}''$ (that is, the dual ideal generated by their set union) is likewise their meet in $\mathfrak{L}$. Then for $c \in A \cap B$ we have that $a \cap b \leq c$ by Dilworth ([5], Definition 2.1).[10] Thus again $a \cap b$ generates $A \cap B$. *Case* 3: $A \in \mathfrak{J}'$, $B \in \mathfrak{J}''$. Then by Theorem 4′, $A \cap B = A \cap \mathfrak{J}'(B)$. Now $A \cap \mathfrak{J}'(B)$ is generated by $a \cap b$, for $a \cap b \in A \cap \mathfrak{J}'(B)$, and if $c \in A \cap \mathfrak{J}'(B)$ then $c \leq a$, $c \leq b$, so $c \leq a \cap b$.

THEOREM 5′. *For any splitting of a given lattice* L, $\mathfrak{L} = \mathfrak{J}' + \mathfrak{J}''$ *is modular (distributive) if and only if conditions* (I) *and* (II) *of Theorem* 5 *hold for* $\mathfrak{E}(J')$ *and* $\mathfrak{E}(J'')$ *in* $\mathfrak{J}'$ *and* $\mathfrak{J}''$.

Proof. From Theorem 4″ it follows that $\mathfrak{L}$ is split into $\mathfrak{J}'$ and $\mathfrak{J}''$. That the splitting is principal is shown by the proof of Theorem 4′. Thus Theorem 5 is applicable.

COROLLARY 6.3. *If the lattice* $\mathfrak{L}$ *obtained from a given lattice* L *is modular (distributive) then so is* L.

Proof. By the isomorphism stated in Theorem 4″.

---

[10] This condition was discovered independently by the author (cf. footnote 4) but not published.

NOTE 6.3. In this connection we call attention to a result of Dilworth ([5] Lemma 2.2) in a slightly different situation: The lattice of all ideals of a fixed lattice L is modular (distributive) if and only if L is modular (distributive).

7. Chippings. Each splinter of a lattice L is again a (sub-)lattice under the original inclusion relation. Thus a splinter may itself be splittable, and so on. This leads us to consider the result of successive splittings S_1, S_2, etc. We view the S_i as operations on the lattices which they split and indirectly on L, and call the resulting partitions of the elements of L **chippings**, denoted $\boldsymbol{C_m = (L, S_1, S_2, \cdots, S_m)}$. We suppose m finite though perhaps unbounded.

A partition class c of the chipping C_m is called a **chip** of C_m, and we write $\boldsymbol{c \in C_m}$. Each chip is a convex sublattice of L, but the converse is not true; cf. **3**. The operations S_i are obviously permutable if the chips which they split have no elements in common; thus the sequence of splittings—the **path**—by which a chipping can be obtained is not unique. The difference can be even greater, as may be seen by considering $FL(2)$. However, the number of chips of C_m (hence the number of splittings by which it is obtained) is invariant.

We shall say $\boldsymbol{C' > C''}$ if and only if C'' can be obtained from C' by further splitting. This direction of the inclusion sign is contrary to the concept of C'' as a successor of C', but accords with customary usage for partitions (cf. [3] p. 6; [6]). Obviously, for chips $c' \in C'$ and $c'' \in C''$, the set intersection $c'c''$ is a convex sublattice of L; if $C' \geq C''$, then $c'c''$ is c'' or void. If $c \in C_m = (L, S_1, \cdots, S_k, \cdots, S_m)$, and k is given, we denote by $g_c(C_k, S_{k+1}, \cdots, S_m)$ the smallest number of the given splittings by which one can obtain from C_k a chipping with c as a chip. We may say that c is of the g_c-th generation after C_k. We write

$$g(C_k, S_{k+1}, \cdots, S_m) = \max_{c \in C_m} g_c(C_k, S_{k+1}, \cdots, S_m).$$

By definition, $g_c(C_k, S_{k+1}, \cdots, S_m)$ and $g(C_k, S_{k+1}, \cdots, S_m)$ are independent of the path by which C_k is obtained, but they may depend on $S_{k+1}, \cdots, S_m$. Obviously

LEMMA 7.1. $g(C_k, \cdots, S_{m+j}) \leq g(C_k, \cdots, S_m) + g(C_m, \cdots, S_{m+j})$.

LEMMA 7.2. *Given* $C' = (L, S_1, \cdots, S_k, S'_{k+1}, \cdots, S'_m)$ *and* $C'' = (L, S_1, \cdots, S_k, S''_{k+1}, \cdots, S''_n)$, *then there exists* $C''' = (L, \cdots, S'_m, S'_{m+1}, \cdots, S'_p) = (L, \cdots, S''_n, S''_{n+1}, \cdots, S''_p)$ *such that* (1) *the chips of* C''' *are the non-void set intersections* $c'c''$ *where* $c' \in C'$ *and* $c'' \in C''$, *and*

(2) $g(C', \cdots, S'_p) \leq g(C_k, \cdots, S''_n)$ *and* $g(C'', \cdots, S''_p) \leq g(C_k, \cdots, S'_m)$.

We denote C''', which is uniquely determined by (1), by $\boldsymbol{C' \cap C''}$. Condition (2) serves for induction.

Proof. The intended paths being obvious, we abbreviate $g(C', \cdots, S'_p)$ to $g(C'/C'_p)$. The lemma is easily verified if $g(C_k/C')$ or $g(C_k/C'')$ is zero or both are 1. Then we proceed by induction on $z = g(C_k/C') + g(C_k/C'')$. Suppose the lemma is true for $z < z_0$ and that we have C' and C'' with $z = z_0$. If this has not already been taken care of, then say $g(C_k/C'') = p > 1$. It is easily seen that there is another path to C'', say $C'' = (L, \cdots, S_k, S^*_{k+1}, \cdots, S^*_n)$, such that for some j, $g(C_k/C^*_{k+j}) = 1$ and $g(C^*_{k+j}/C'') = p - 1$. Denote C^*_{k+j} by C^*. Now

$$g(C_k/C') + g(C_k/C^*) = g(C_k/C') + 1 < z_0,$$

so by induction there exists $C' \cap C^*$, and $g(C^*/C' \cap C^*) \leq g(C_k/C')$. Then

$$g(C^*/C'') + g(C^*/C' \cap C^*) \leq p - 1 + g(C_k/C') < z_0,$$

so again by induction there exists $(C' \cap C^*) \cap C''$, which we call C'''. One can verify that (1) and (2) hold. Q. E. D. We may remark that it would not suffice to make an induction on the length of the path, for this does not obey (2).

THEOREM 8. *The set of all chippings of a lattice is a lattice.*

Proof. We use the inclusion relation defined above. The $C' \cap C''$ of Lemma 7. 2 is a lower bound to C' and C'' by construction; that it is greatest follows from (1) of that lemma. As for least upper bound, we observe that the ascending chain condition holds and that the set of chippings has a unit, namely (L). Then least upper bounds exist by an argument like the proof of Theorem 2. 2 of [3] (p. 17).

Condition (1) of Lemma 7. 2 shows that $C' \cap C''$ is just the common refinement of C' and C'' as partitions; however, examples (using the convexity of chips) show that the join of C' and C'' as partitions need not be a chipping. Thus the lattice of chippings is not a sublattice of the lattice of all partitions. In fact, though the latter is a matroid lattice (Birkhoff [3] p. 67) the former need not be, as is seen in the case of $FL(2)$.

From Theorem 8 we see that the unsplittable chips of a lattice are uniquely determined. From **3** it follows that any finite distributive lattice has a chipping whose chips consist of single elements. Also, we might consider only principal splittings. The above arguments (e. g., Theorem 8) can equally well be carried out under this restriction. In this case we can define the edge of a chipping as the collection of all the elements of the edges of the splittings involved, with proper multiplicity. By examining simple cases as in the proof of Lemma 7. 2, and induction as in the proof of the Jordan-Hölder Theorem, it can be

shown that the edge of a chipping C is independent of the path by which C is reached.

8. Unsolved problems. (1) Characterize lattices of partitions[11] ([3] p. 146). This would be partly done if one could show that $FL(n)$ could be represented as such a lattice. (2) Characterize $C' \cup C''$ (**7**). (3) Can Stone's and Birkhoff's work (cited in **3**) on splitting distributive lattices be generalized? We know from **3** that not every modular lattice can be split. Ward and Dilworth have generalized this work in a slightly different direction ([9] p. 332). (4) What further properties has the lattice of all chippings of a fixed lattice?

HARVARD UNIVERSITY
AND
THE UNIVERSITY OF PENNSYLVANIA.

BIBLIOGRAPHY

[1] G. Birkhoff, "On the combination of subalgebras," *Proceedings of the Cambridge Philosophical Society*, Vol. 29 (1933), pp. 441-464.
[2] ———, "On the structure of abstract algebras," *ibid.*, Vol. 31 (1935), pp. 433-454.
[3] ———, "Lattice Theory," *American Mathematical Society Colloquium Publications* XXV, New York (1940).
[4] ———, "Rings of sets," *Duke Mathematical Journal*, Vol. 3 (1937), pp. 443-454.
[5] R. P. Dilworth, "Ideals in Birkhoff lattices," *Transactions of the American Mathematical Society*, Vol. 49 (1941), pp. 325-353.
[6] P. Dubreil and M.-L. Dubreil-Jacotin, "Théorie algébrique des relations d'équivalence," *Jour. de Math.*, Vol. 18 (1939), pp. 63-95.
[7] M. H. Stone, "The theory of representations for Boolean algebras," *Transactions of the American Mathematical Society*, Vol. 40 (1936), pp. 37-111.
[8] ———, "Topological representations of distributive lattices and Brouwerian logics," *Casopis pro Pestovani Mat. a Fys.*, Vol. 67 (1937), pp. 1-25.
[9] M. Ward and R. P. Dilworth, "Evaluations over residuated structures," *Annals of Mathematics* (2), Vol. 40 (1939), pp. 328-337.
[10] P. M. Whitman, "Free lattices," *Annals of Mathematics* (2), Vol. 42 (1941), pp. 325-330.
[11] ———, "Free lattices II," *ibid.*, Vol. 43 (1942), pp. 104-115.

[11] Added in proof: In this connection see O. Ore, "Theory of equivalence relations," *Duke Mathematical Journal*, Vol. 9 (1942), pp. 573-628.